

US Patent No. 5763188 Page 9

Query Match 3.5%; Score 31.4; DB 4; Length 1119;
Best Local Similarity 61.7%; Pred. No. 3.1;
Matches 50; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

28 ccccttcgcaacccgcatgacacccgtgttgcacccgtctcaaatgcttcctc 87
187 cccgtatgctgctgcgcatacaacgtgctgcgcaccccttcttcgacatcaag 246

88 ctgcgccttcctctgtctgcac 108
247 ctgtccctctgctgcacatgcc 267

RESULT 18
US-08-987-943-1
Sequence 1, Application US/08987943
Patent No. 6210967
GENERAL INFORMATION:
APPLICANT: Bard, Jonathan A.
TITLE OF INVENTION: DNA Encoding A Mammalian LPA Receptor and Uses Thereof
FILE REFERENCE: 55183
CURRENT APPLICATION NUMBER: US/08/987,943
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentln Ver. 2.0 - beta
SEQ ID NO 1
LENGTH: 1755
TYPE: DNA
ORGANISM: Homo sapiens
US-08-987-943-1

Query Match 3.5%; Score 31.4; DB 4; Length 1755;
Best Local Similarity 61.7%; Pred. No. 4;
Matches 50; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

28 ccccttcgcaacccgcatgacacccgtgttgcacccgtctcaaatgcttcctc 87
536 ccggtatgctgctgcgcatacaacgtgctgcgcaccccttcttcgacatcaag 595

88 ctgcgccttcctctgtctgcac 108
536 ctgtccctctgctgcacatgcc 918

RESULT 19

Sequence 15, Application US/08920812
Patent No. 5763188
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,812
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 5247 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Pseudomonas aeruginosa
STRAIN: Clinical Isolate P2-17
US-08-920-812-15

Query Match 3.5%; Score 31.4; DB 1; Length 5247;
Best Local Similarity 64.4%; Pred. No. 7.2;
Matches 47; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

42 cgcacatgacacccgtgttgcacccgtgttgcacccgtcttcctcctcgtcgcgttcctg 101
698 ccgcattgacttgcgccttcacacccgtgttgcacccgtcttcctcctcgtcgcgttcctg 639

102 tctgcacacgtg 114
638 ccgcacacacgtg 626

RESULT 20

Sequence 15, Application US/08920827
Patent No. 5770375
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,827
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448

100

REFERENCE/DOCKET NUMBER: 17796-002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 877-0900
TELEFAX: (415) 877-8370
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2070 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US91-07290-6

[illegible]

Db 678 GCCTCTACAGAGGGGACACCGACAGAAAGAACACCAACACCATCATCAACAC 737
 2Y 444 catgtacagccgcgcgaataatcaagcgatagacaaatcatgtcagggcgggtgcg 503
 Db 738 Cgtggccctcgtgacgtcaacgtcgtcgtgggagtggaaggttcgagggcggtggcg 797
 QY 504 cggcnaagggcaaacgc 545
 Db 798 CGCACCGAGGACATCAAGACCGACGCTCACCGAGAGCTGAA 839

Result 20

US-07-642734C-3
 Sequence 3, Application US/07642734C
 Patent No. 5824513

GENERAL INFORMATION:

APPLICANT: Katz, L.
 APPLICANT: Donadio, S.
 APPLICANT: Mcalpine, J. B.
 TITLE OF INVENTION: Recombinant DNA Method for Producing
 TITLE OF INVENTION: Erythromycin Analogs
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Edward H. Gorman
 STREET: Abbott Laboratories D377/AP6D-2 One Abbott
 CITY: Park Rd
 STATE: IL
 COUNTRY: US

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/642,734C
 FILING DATE: 17-JAN-91
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Danckerts, Andreas M
 REGISTRATION NUMBER: 32652
 REFERENCE/DOCKET NUMBER: 4957-05.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 708-937-9396
 TELEFAX: 708-938-2623
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20235 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Saccharopolyspora erythraea
 STRAIN: NRRL 238

FEATURE:

NAME/KEY: CDS
 LOCATION: 19..10722
 OTHER INFORMATION: /codon_start=19
 OTHER INFORMATION: /function="gene eryA"
 OTHER INFORMATION: /product="eryA ORF2 encoding modules 3 & 4 for
 OTHER INFORMATION: 6-deoxyerythronolide B"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 19..4470
 OTHER INFORMATION: /function="approximate span of
 OTHER INFORMATION: module 3"
 FEATURE:
 NAME/KEY: misc_feature

LOCATION: 97..1482

OTHER INFORMATION: /function="approximate span of
 OTHER INFORMATION: beta-ketoacyl ACP synthase of module 3"

NAME/KEY: misc_feature

LOCATION: 1693..2670
 OTHER INFORMATION: /function="approximate span of
 OTHER INFORMATION: acyltransferase domain module 3"

NAME/KEY: misc_feature

LOCATION: 3406..3921
 OTHER INFORMATION: /function="approximate span of
 OTHER INFORMATION: beta-ketoreductase domain of module 3"

NAME/KEY: misc_feature

LOCATION: 4171..4428
 OTHER INFORMATION: /function="approximate span of
 OTHER INFORMATION: acyl carrier domain of module 3"

NAME/KEY: misc_feature

LOCATION: 4471..5072
 OTHER INFORMATION: /function="approximate span of
 OTHER INFORMATION: module 4"

NAME/KEY: misc_feature

LOCATION: 4471..5847
 OTHER INFORMATION: /function="approximate span of
 OTHER INFORMATION: beta-ketoacylacyltransferase domain of module 4"

NAME/KEY: misc_feature

LOCATION: 6054..7026
 OTHER INFORMATION: /function="approximate span of
 OTHER INFORMATION: acyltransferase domain of module 4"

NAME/KEY: misc_feature

LOCATION: 7165..9216
 OTHER INFORMATION: /function="approximate span of
 OTHER INFORMATION: dehydratase and enoylreductase domains m"

NAME/KEY: misc_feature

LOCATION: 9433..9984
 OTHER INFORMATION: /function="approximate span of
 OTHER INFORMATION: beta-ketoreductase of module 4"

NAME/KEY: misc_feature

LOCATION: 10225..10483
 OTHER INFORMATION: /function="approximate span of
 OTHER INFORMATION: acyl carrier domain of module 4"

NAME/KEY: misc_feature

LOCATION: 10723..15165
 OTHER INFORMATION: /function="approximate span of
 OTHER INFORMATION: module 5"

NAME/KEY: misc_feature

LOCATION: 10831..12174
 OTHER INFORMATION: /function="approximate span of
 OTHER INFORMATION: beta-ketoacylacyltransferase domain of module 5"

NAME/KEY: misc_feature

LOCATION: 12379..13350
 OTHER INFORMATION: /function="approximate span of
 OTHER INFORMATION: acyltransferase domain of module 5"

NAME/KEY: misc_feature

LOCATION: 14062..14610

OY 3730 aacctggcaggcgccgctgcgcgcatccgtgttttcgcacaacacgcgaacacaccc 3789
 Db 3600 GCCTTACCTAATGACGAATATGGGGCAGTTTCTTCGCAAGCCCTTCAATATACCTT 3689
 OY 3790 gagcagcgcatcgcaactcgcagcggcttgcgccacgcgcgcttttcggcaatgcgc 3849
 Db 3660 GATGAAACAGGTTAAATAATCACGGACATTTAACGATGATGTCGGGTTTGGCCCAATCA 3719
 OY 3850 atgcacaggttcacatcgcgcacgcagcgcgcggtttagcagcgacgctttca 3909
 Db 3720 TGGGGCATTACCAATTGGTGTAAACGTTGGGACGGGAATCAGTCCGAGTAAATGGCT 3779
 OY 3910 gagcgatcggaagcaaatccgcgcgcgcgcgcgtgcattcaggatcagcgacgatac 3969
 Db 3780 GAAGACAAACCCCAAAAATTCATCCAAAAACCATTAATATAGGCTCAATATCAAGTAT 3839
 OY 3970 cgcgcgcggtttcgcggtattcgcatcgcgacgcgcgcacatcgcgcgaacgcgtattcgtc 4029
 Db 3840 CAGTTCGGTTTAAAGGCAATATGGCATTCAGGCTTTATTTGGAGTTAATCGCTATTTTAT 3899
 OY 4030 caaaaacgagattaccgcgcacgaagaagttcaatcgcaccccccgcgcgttcgatac 4089
 Db 3900 GAACGTAAATTTATCAATCTGAGTACATGACGATGACGATGAAAAACGCTTAGCTTGCATTTAT 3959
 OY 4090 cgcctaccgcgcggcattagaggatattcattcaaacgcggcgcaaacatctcatc 4149
 Db 3960 CGCTATATATCTGGCATTCAGATGTATTTATACATTATACCCGACGATTAATATACGCTT 4019
 OY 4150 acgcctattctggccgtctcctatccgctgcgcgttcggcgaaagtcgcgaacgcgctc 4209
 Db 4020 AAGCTTATTTCTTCTGTCATTTGTTGATGTTTCAAAACGCTACGTCMAAACACGATA 4079
 OY 4210 aatcgccggtatttgctcgaattcgcgaataaccgcgcagtgcggaatggggcgtaac 4269
 Db 4080 AATGTCAAGGTTGTCACAAACCATTTGACGTTATTTGCCAAAAAGAAAGTGCATTTAAG 4139
 OY 4270 gcccaatcaaaagtttcaacgcgtctccctccacgcgcgcgcgcgaaaagccccgcactg 4329
 Db 4140 GCAGAAATTTTACATTTCCAAATTTCCGCTTTATTCATAAATTCGAAGTTCAACAACCTO 4199
 OY 4330 gaagcgcaacacagcgcggcgcacatlaattagctacgcctgttaa 4374
 Db 4200 GGGTACGCGTAAATGAGGCGTAAATGGCTATCGTTGGTAA 4244

RESULT 15
 NLIGA3327 578 bp DNA linear
 LOCUS NLIGA3327
 DEFINITION Neisseria lactamica Iga protease homologue gene, isolate
 ACCESSION AJ001740
 VERSION AJ001740.1 GI:3152399
 KEYWORDS Iga protease.
 SOURCE Neisseria lactamica.
 ORGANISM Neisseria lactamica
 Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.

REFERENCE
 AUTHORS Sayers, J.R.
 TITLE Direct Submission
 JOURNAL Submitted (20-OCT-1997) Sayers J.R., Division of Molecular
 MEDICINE, University of Sheffield, Royal Hallamshire Hospital,
 Sheffield, S. Yorks, S10 2UF, ENGLAND
 REFERENCE
 AUTHORS Sayers, J.R., Viltowski, S. and Read, R.
 TITLE Strains of Neisseria lactamica carry an Iga1 protease gene
 homologue and some express Iga1 protease activity
 JOURNAL Unpublished
 FEATURES
 location/Qualifiers
 1..578
 /organism="Neisseria lactamica"
 /isolate="NL3327"
 /db_xref="taxon:486"

CDS

```
<1..>576
/codon_start=1
/transl_table=11
/product="Iga protease"
/protein_id="CAA04965.1"
/db_xref="GI:3152400"
/db_xref="SPTREMBL:O69746"
/translation="GKPAVGAADIEVYKKGELVSKMTKAPMIDFVSRNGVAALV
GQYIVSAHNGGYNVDFGAGSNPDHRSFOIVKRNKYPDNDHPYNGDYMPL
HFTVDAEPVEMTGDMRGNTYSDEKRYEPRIKSGHHYWRYYDDDKHGLDLSYSGAWLI
GGTHMOGNGNNGVSVLSGDVHRANDYGPMP1"
```

```
repeat_region
repeat_region 139..148
BASE COUNT 163 a 125 c 156 g 134 t
ORIGIN
```

```
Query Match 7.3%; Score 318.8; DB 1; Length 578;
Best Local Similarity 83.3%; Pred. No. 4.1e-61;
Matches 375; Conservative 0; Mismatches 72; Indels 3; Gaps 1;
```

```
OY 184 ggcaagttgcagtcgagcgaggaagabattgaggttaacacaaagaggtgtc 243
Db 1 GGCAAGTTGTCAGTCGAGCGGAGGAAGATTTGAGCTTACACAAAAGGCGGTGTC 60
OY 244 ggcaatcaatgacaaagcccgatgatttctctgtgtgtcgcgttaacgagcgt 303
Db 61 GGCAATCAATGACAAAGCCCGATGATTCTTTCGTGTGTGTCGCTAACGCGGTG 120
OY 304 ggcgcatgtgtgagcgatcaatattgtgagctggaacaaagcggtacataaac 363
Db 121 GCGCATGTGTGAGCGGATCAATATTGTGACGCTGACATACGCGGCTATACAAAT 180
OY 364 gtcgatttgcgaggaagaaatcccgatcaaatcttcttactataaatgtg 423
Db 181 GTTGATTTGTCGTCGAGGAGCAATCCGATCAGCACCGCTTTCTTACAAATGTG 240
OY 424 aaacggaataataagcagaggaacaaagccatccatctatgcgagctatcatg 483
Db 241 AAAAGAAATCACTCAAAACCTGACATTTCA--CACCTTACAAAGCGGCTATAC 297
OY 484 cgcggttgcgaatattgtcagagtcgagaaacgttttaaatgcccagttatgat 543
Db 298 CCGCGCTGCAATAGTCGATGACGATGACGAACTGTGCAAAATGACGGGTGACATG 357
OY 544 gggcggaataatcagatcaaatcaatcaatccctgaccgtgtcgtatgtggagcag 603
Db 358 GGGATACCTTATCCGATTAAGAAATATCCGAGCGGTGCGCATCGCATGAGAC 417
OY 604 caatctgcgagtcgagatgagagcc 633
Db 418 CACTTTGCGCTTATGATGATGACAAACAC 447
```

RESULT 16

```
NTIGA3354 584 bp DNA linear BCT 21-MAY-1998
DEFINITION Neisseria lactamica Iga protease homologue gene, isolate NL3354.
ACCESSION AJ001737
VERSION AJ001737.1 GI:3152397
KEYWORDS Iga1 protease.
SOURCE Neisseria lactamica.
ORGANISM Neisseria lactamica; beta subunit; Neisseriaceae;
```

```
REFERENCE 1 (bases 1 to 584)
AUTHORS Sayers, J.R.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1997) Sayers J.R., Division of Molecular
Medicine, University of Sheffield, Royal Hallamshire Hospital,
Sheffield, S. Yorks, S10 2JF, ENGLAND
REFERENCE 2 (bases 1 to 584)
```

```
AUTHORS Sayers, J.R., Witowski, S. and Read, R.
TITLE Strains of Neisseria lactamica carry an Iga1 protease gene
homologue and some express Iga1 protease activity
JOURNAL Unpublished
FEATURES Location/Qualifiers
Source 1..584
/organism="Neisseria lactamica"
/isolate="NL3354"
/db_xref="taxon:486"
<1..>582
/codon_start=1
/transl_table=11
/product="Iga1 protease"
/protein_id="CAA04964.1"
/db_xref="GI:3152398"
/db_xref="SPTREMBL:O69747"
/translation="FAVGAKDIEVYKKGELVSKMTKAPMIDFVSRNGVAALVD
OYIVSAHNGGYNVDFGAGSNPDHRSFOIVKRNKYPDNDHPYNGDYMPL
HFTVDAEPVEMTGDMRGNTYSDEKRYEPRIKSGHHYWRYYDDDKHGLDLSYSGAWLI
GGTHMOGNGNNGVSVLSGDVHRANDYGPMP1A"
```

CDS

```
repeat_region
repeat_region 144..148
BASE COUNT 163 a 127 c 159 g 185 t
ORIGIN
```

```
Query Match 7.2%; Score 313.6; DB 1; Length 584;
Best Local Similarity 81.6%; Pred. No. 6e-60;
Matches 368; Conservative 0; Mismatches 69; Indels 3; Gaps 1;
```

```
OY 190 ttgcagtcgagcgaggaagattgattgattacacaaagaggtgtgtgcgcaaa 249
Db 1 TTGCAGTCGAGCGGAGGAAGATTGATTGATTACACAAAAGGAGTTGTGTCGCAAA 60
OY 250 tcaatgcaaaagcccgatgatttctctgtgtgtcgcgttaacgagcgtgtgcga 309
Db 61 TCAATGACAAAGCCCGATGATTCTTTCGTGTGTGTCGCTAACGCGCTGAGCA 120
OY 310 ttgctggcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 369
Db 121 TTGCTGGCGGATCAATATTGTGAGCGTGCACATACGCGGCTATACAAATGTGAT 180
OY 370 ttgctggcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 429
Db 181 TTGCTGGCGGATCAATATTGTGAGCGTGCACATACGCGGCTATACAAATGTGAT 240
OY 430 aataatgaaagcaggaactaaagccatccctatgagcgagtcgagtcgagtc 489
Db 241 AATAATGAAAGCAGGAACTGAAATTTCA--CACCTTACAAAGCGGCTATAC 297
OY 490 ttgcatgaattgtcacagatgacgaacctgttgaattgacagtcataatgagcg 549
Db 298 CTGCTAATGTTGTAACGATGACGAACTGTGAAATGACGGGTGACATGAGCGGAAT 357
OY 550 aatctatcgatcaaatcaatcaatccctgaccgtgtcgtatgtggcagcgcaat 609
Db 358 AACTATCGATCAAAATCAATCAATCCGATGAGCGGTGCGCATCGCTGAGACAC 417
OY 610 tggcgatctgagatgagatga 629
Db 418 TGGCGTATGATGATGAGCA 437
```

RESULT 17

```
NTIGA823 572 bp DNA linear BCT 21-MAY-1998
DEFINITION Neisseria lactamica Iga protease homologue gene, isolate NL823.
ACCESSION AJ001737
VERSION AJ001737.1 GI:3152389
KEYWORDS Iga protease.
SOURCE Neisseria lactamica.
ORGANISM Neisseria lactamica
```

```
REFERENCE 1 (bases 1 to 572)
AUTHORS Sayers, J.R.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1997) Sayers J.R., Division of Molecular
Medicine, University of Sheffield, Royal Hallamshire Hospital,
Sheffield, S. Yorks, S10 2JF, ENGLAND
REFERENCE 2 (bases 1 to 572)
```


2264 TGAATATGCAAGCGGAGTGTAAAGTTTACCAAACTTAATGAGATGTCATTAAAC 2323
2349 CGGAGGACGCGCCTAAGCTAAGCTAAGCGGAGGACCAACCAAGCGCAGCTCGT 2408
2324 AAATCAACAGCC----- 2335
2409 GGGCAATGCCAAGCAATTAACAGCCATTAACGCGGAGCAACATCGCTCGA 2469
2336 ----- 2435
2469 CAATGCTCAATTAATCAAGCAACAGCGGAGGAGCGGAGCGCTCGCTCGA 2528
2336 -----ATTACATTTAGACCAATGCGCAACCAATGCGCAATTTGATTTGCA 2386
2529 CAAGCTAAGGAGCAAGCAAGCTAAGCGGAGGAGCGGAGCGGAGCGGAGGAG 2588
2387 CAATTAATGCAAGCGGAGTATGCAAACTTACAGGATGATGATTTAGCGATTC 2446
2589 GAGCAGTATCACTTTGAGAAAGCGCTTACCGGAGAAATCAGCGGAGGAGGATC 2648
2447 AGCTCAATTTCTTTAAAGCAAGCATTTTTCGACCAAAATTCAGCGGAGGAGG 2506
2649 GAGCAGTATCACTTTGAGAAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2708
2507 AACAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2566
2709 AAGCCTTGAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2768
2567 AACGTAATTAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2609
2769 GCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2828
2610 -----GCTACGTCACAAACAAATACGCG-----TCGCGCTTCAATTAAGAGCGG 2553
2829 TACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2888
2654 AACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2713
2889 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2948
2714 TGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2773
2949 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3008
2774 ATTATCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2833
3009 CGTAAAGTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3068
2834 CGAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2893
3069 TCTTAATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3128
2894 GCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2953
3129 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3188
2954 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3013
3189 ACTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3248
3014 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3073
3249 CAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3308
3074 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3133
3309 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3368
3434 TGAATACCTCTGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3193
3369 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3428
3194 TGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3253

3429 CTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3488
3254 TTTGATATCCCTGCTGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3313
3489 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3446
3314 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3373
3547 -----GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3599
3374 ACAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3433
3600 CAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3659
3434 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3493
3660 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3719
3494 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3553
3720 CTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3776
3554 TTTATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3613
3777 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3836
3614 ACGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3673
3837 CAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3896
3674 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3733
3897 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3956
3734 ATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3793
3957 CAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4016
3794 AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3853
4017 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4076
3854 GCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3913
4077 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4136
3914 TCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3973
4137 CATTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4196
3974 CATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4033
4197 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4256
4034 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4093
4257 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 416
4044 GCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4153
4117 TTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4376
4154 TTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4213
4377 CGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4407
4214 TGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG

Sayers, SR

Seed ID no: 65

[illegible]

LOCUS NLGA3327 578 bp DNA linear BCT 21-MAY-1998
DEFINITION *Neisseria lactamica* Iga protease homologue gene, isolate NL3327.
ACCESSION AJ001740
VERSION AJ001740.1 GI:3152399
KEYWORDS Iga protease.
SOURCE *Neisseria lactamica*.
ORGANISM *Neisseria lactamica*.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; *Neisseria*.

REFERENCE
AUTHORS Sayers, J.R.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1997) Sayers J.R., Division of Molecular Medicine, University of Sheffield, Royal Hallamshire Hospital, Sheffield, S. Yorks, S10 2JF, ENGLAND
2 (bases 1 to 578)
Sayers, J.R., Witowski, S. and Read, R.
TITLE Strains of *Neisseria lactamica* carry an Iga1 protease gene homologue and some express Iga1 protease activity
JOURNAL Unpublished
FEATURES Location/Qualifiers

AUTHORS	Meyer,T.F., Rudel,T., Scheuerflug,I., Fischer, Eckhard and Maier,J.
TITLE	NUCLEIC ACID MOLECULES WHICH CODE PROTEINS WHICH MEDIATE THE ADHESION OF NEISSERIA CELLS TO HUMAN CELLS
JOURNAL	PATENT: WO 9711181-A 9 27-MAR-1997; MAX PLANCK GESELLSCHAFT (DE) Other publication DE 19534579 970320.
COMMENT FEATURES	Location/Qualifiers 1..1744
SOURCE	/organism="Neisseria gonorrhoeae" /strain="MS11" /db_xref="taxon:485"
BASE COUNT	560 a 383 c 405 g 396 t
ORIGIN	
Query Match	7.3%; Score 121.8; DB 6; Length 1744;
Best Local Similarity	66.7%; Pred. No. 1e-20; Mismatches 87; Indels 0; Gaps 0;
Matches 174; Conservative 0; Mismatches 87; Indels 0; Gaps 0;	
Db	237 ccgagcaggcgccatgaagaataacgcgtctacattgctccgcttcgcatacagggca 296 284 CGAAGCGACGGGTTGAAGCATTAATCGGTAAAGAACCCATTTCAGGACACGGACA 343 Qy 297 cgaaagtccattcccccttgcacaaccatgcctcacatccgatccgatcatgaagccgtag 356 Db 344 CGAAGTTCACAGCTCCGTTGCATTAATCAATCATTCAAAAAGCACTTCTATTTTCAGCGCGG 403 Qy 357 tcccgttgacgattcagcctttaccgcacatccattggagagataagaacatccgcg 416 Db 404 CGTAGACGGGGCTTTTACGTTTACCATTCAACTTCATCGCACAGGCTCGGAATAATACATCCGC 463 Qy 417 cgacgcctatgacggcgccacaggcgcgcgctatccgccctcccaaaagcgcgaggatat 476 Db 464 AGACGGATATGACGGGCTCTCAAGCGCGGGTATTCGGAAACCAAGGCGGACGAGTAT 523 Qy 477 ataccagtaacataaagg 497 Db 524 ATACGCTACCATTAATCAAAG 544
LOCUS	A61821 3287 bp DNA linear PAT 09-MAR-1996
DEFINITION	Sequence 1 from Patent WO9711181.
ACCESSION	A61821
VERSION	A61821.1 GI:3715995
KEYWORDS	
SOURCE	Neisseria gonorrhoeae.
ORGANISM	Neisseria gonorrhoeae Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria
REFERENCE	1 (bases 1 to 3287)
AUTHORS	Meyer,T.F., Rudel,T., Scheuerflug,I., Fischer, Eckhard and Maier,J.
TITLE	NUCLEIC ACID MOLECULES WHICH CODE PROTEINS WHICH MEDIATE THE ADHESION OF NEISSERIA CELLS TO HUMAN CELLS
JOURNAL	PATENT: WO 9711181-A 1 27-MAR-1997; MAX PLANCK GESELLSCHAFT (DE) Other publication DE 19534579 970320. Location/Qualifiers 1..3287
COMMENT FEATURES	/organism="Neisseria gonorrhoeae" /strain="MS11" /db_xref="taxon:485" clone_1jb-"(GENOMISCHE BIBLIOTHEK IN PBA)" clone_1lb-"(B) CLONE(E): H1967/PBS25"
BASE COUNT	1017 a 740 c 752 g 778 t
ORIGIN	
Query Match	7.3%; Score 121.8; DB 6; Length 3287;
Best Local Similarity	66.7%; Pred. NO. 1.2e-20;

Meyer et al.

SEQ ID NO. 463

Db 425 ACGGATATGAGGCGCTTAAAGCGCGCTTAAAGGATCCAGCAAGGAGGATATAT 484
Oy 479 aaagctacacataaag 497
Db 485 ACGGCTACCAATATCAAGG 503

RESULT 30

DEFINITION Sequence 9 from Patent WO9711181.
ACCESSION A61829 1744 bp DNA linear PAT 09-MAR-1998
VERSION A61829
KEYWORDS A61829.1 GI:3715998
SOURCE Neisseria gonorrhoeae.
ORGANISM Neisseria gonorrhoeae
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;

REFERENCE 1 (bases 1 to 1744)
AUTHORS Meyer, T.F., Rudel, T., Scheuerpflug, I., Fischer, Eckhard and Maier, J.

TITLE NUCLEIC ACID MOLECULES WHICH CODE PROTEINS WHICH MEDIATE THE ADHESION OF NEISSERIA CELLS TO HUMAN CELLS
JOURNAL Patent: WO 9711181-A 9 27-MAR-1997;
MAX PLANCK GESSELLSCHAFT (DE)
COMMENT Other publication DE 19534579 970320.
FEATURES Location/Qualifiers
source 1. 1744

BASE COUNT 560 a 383 c 405 g 396 t
ORIGIN /organism="Neisseria gonorrhoeae"
/strain="MS11"
/db_xref="taxon:485"

Query Match 6.3%; Score 119.8; DB 6; Length 1744;
Best Local Similarity 66.4%; Pred. No. 1.2e-17;
Matches 172; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Oy 239 aacagcgccgcttgaagaaatcgcctacatgttcgctttccgatacagggacaca 298
Db 286 AACGACAGGCGTTTACCGCTTATGAAACCCATTTCAGACGACGACGACG 345
Oy 299 aatccatcgccttcgacacacatccatccatccgattcgaacgacgagtagtc 358
Db 346 AAGTACACAGCGCTTCATATCATATCAAAAGCACTTCATTTACGCGCGCG 405
Oy 359 ccgttgagagatcgccttaccgcatcattggagagataagacacatccgcgcg 418
Db 406 TAGACGGCGGTTTACCGCTTACCACTTCATCGGACAGGCGTGGAAATACATCCCGACG 465
Oy 419 acgctatgacgagcgaagcgagcgatccgcctcccaaaagcgagcgagtagtc 478
Db 466 ACGGATATGAGGCGCGCTTAAAGCGCGCTTAAAGGATCCAGCAAGGAGGATATAT 525
Oy 479 aaagctacacataaag 497
Db 526 ACGGCTACCAATATCAAGG 544

RESULT 31

LOCUS A61821 3287 bp DNA linear PAT 09-MAR-1998
DEFINITION Sequence 1 from Patent WO9711181.
ACCESSION A61821
VERSION A61821.1 GI:3715995
KEYWORDS A61821.1
SOURCE Neisseria gonorrhoeae.
ORGANISM Neisseria gonorrhoeae
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;

REFERENCE 1 (bases 1 to 3287)

AUTHORS Meyer, T.F., Rudel, T., Scheuerpflug, I., Fischer, Eckhard and Maier, J.
TITLE NUCLEIC ACID MOLECULES WHICH CODE PROTEINS WHICH MEDIATE THE ADHESION OF NEISSERIA CELLS TO HUMAN CELLS
JOURNAL Patent: WO 9711181-A 1 27-MAR-1997;
MAX PLANCK GESSELLSCHAFT (DE)
COMMENT Other publication DE 19534579 970320.
FEATURES Location/Qualifiers
source 1. 3287

BASE COUNT 1017 a 740 c 752 g 778 t
ORIGIN /organism="Neisseria gonorrhoeae"
/strain="MS11"
/db_xref="taxon:485"
/clone_lib="(B) CLON(E): H1967/PBS25"
/clone_lib="(B) CLON(E): H1967/PBS25"

Query Match 6.3%; Score 119.8; DB 6; Length 3287;
Best Local Similarity 66.4%; Pred. No. 1.3e-17;
Matches 172; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Oy 239 aacagcgccgcttgaagaaatcgcctacatgttcgctttccgatacagggacaca 298
Db 1829 AACGACAGGCGTTTACCGCTTATGAAACCCATTTCAGACGACGACGACG 1888
Oy 299 aatccatcgccttcgacacacatccatccatccgattcgaacgacgagtagtc 358
Db 1889 AAGTACACAGCGCTTCATATCATATCAAAAGCACTTCATTTACGCGCGCG 1948
Oy 359 ccgttgagagatcgccttaccgcatcattggagagataagacacatccgcgcg 418
Db 1949 TAGACGGCGGTTTACCGCTTACCACTTCATCGGACAGGCGTGGAAATACATCCCGACG 2008
Oy 419 acgctatgacgagcgaagcgagcgatccgcctcccaaaagcgagcgagtagtc 478
Db 2009 ACGGATATGAGGCGCGCTTAAAGCGCGCTTAAAGGATCCAGCAAGGAGGAGGATATAT 2068
Oy 479 aaagctacacataaag 497
Db 2069 ACGGCTACCAATATCAAGG 2087

RESULT 32

LOCUS AF142582 3294 bp DNA linear BCT 18-MAY-1999
DEFINITION Neisseria gonorrhoeae maf gene cluster, complete sequence.
ACCESSION AF142582
VERSION AF142582.1 GI:4838552
KEYWORDS Neisseria gonorrhoeae.
SOURCE Neisseria gonorrhoeae
ORGANISM Neisseria gonorrhoeae
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 3294)
AUTHORS Eickelbarger, S., Meyer, T.F., Fischer, E., Maier, J., Manning, P.A., Rudel, T., Scheuerpflug, I., Schulz, E. and Schwan, E.T.
TITLE Direct Submission
JOURNAL Submitted (13-APR-1999) Molekulare Biologie MPI fuer Infektionsbiologie, Monbijoustr. 2, Berlin 10117, Germany
FEATURES Location/Qualifiers
source 1. 3294

organism="Neisseria gonorrhoeae"
strain="MS11"
db_xref="taxon:485"
gene 143..457
/gene="mafI"
143..457
/gene="mafI"
/codon_start=1
/transl_table=1
/product="MafI"
/protein_id="AAD31037.1"

Victoria 3052, Australia
Location/Qualifiers
1. 1356
/organism="Porphyromonas gingivalis"
/strain="W50"
/db_xref="taxon:837"
1. 1356
/codon_start=1
/transl_table=11
/product="50 kD antigen pgl"
/protein="14" "AAD33930.1"
/db_xref="GI:4929280"
/translation="NAKVIKTKKGLALNKGKPLPMLAEPASPTVAVVPDPEVY
PKVTARPDVAVAGSALMHKAYPEMKFSPVSEVIAVNRGAKRVLSIEVRDGLN
EYESPPVDPSSALSAOIKELISSMGFIPKORPYDIATPIDIAEIVITNTAP
LAPDEFIVRGERALQIDAALATKIVKYLKPGSSLSJHNEIVHGPHPAG
NVCYLINHTKPIRNGEYVWLTKLIDILITGKADFTMLAMTGSAAAGYVR
IMPGCNVASFSPGRITIKSHERVLDGNVLTKLEKPELSRDDIIVIPEDDV
DELFGMAVPRIDYMSRAVFSMLOGKNEIYIDAKIKGSEAMTMSNEYDRFPMDI
YPEYLKAIADIDIMEDLGIYEVAPEDFAFCEFDTSKTEIQRVREGLDMLYKEM
N"

CDS

BASE COUNT 338 a 333 c 380 g 305 t
ORIGIN

Query Match 9.1%; Score 122.2; DB 1; Length 1356;
Best Local Similarity 46.0%; Pred. No. 4.2e-19;
Matches 622; Conservative 0; Mismatches 700; Indels 30; Gaps 6;

2 tgatataatcaaaaagcgttaacgtccatcgaggagagcggagcaatcatt 61
11 TTTATAAAACAAAAGGCTGTGCTTACTTATTCGAAAGAAACCGTCCGAGATGC 70
62 atgacggccgcgc---attaccgagtcgcgtgtgtgtgagaaatatacggtatgc 118
71 TGGCCGACCGGCGCAAGTCTTACTTTCGCGGCGGTGCCGACGATTTGAAGTGTGA 130
119 gccctngatgaagtcgaaggaagcgatgcgttcaaaaagcgcaagtgctgttgaag 178
131 TCCCCAAAGGAGCGGCTGTCGCGGGGAGAGTGCCTGCCGCTCACACTGATGACAC 190
179 acaaaagatccggcggt 238
191 ACAAGGATTCGCGAATGAAGTTCACAACTCCGGTAAAGGGCAAGATGATCGGGTGA 250
239 atcgcgagaaagcggtactcagtcgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 298
251 ATGCGGTGCGCAAGCGC---AAGGTGTGAGCATCGAGGTGAACCGGAGGACTGA 304
299 tcgaatcgcaagcgtacgagcgcgcgaagcgttggcaacttaagcgcgangaantngun 358
305 ACGAATACGAGTCTATCCCTGTCGGGGATCCGTCTCTCTGCGCAACAGATCAAG 364
359 gcaatctgataatccggt 418
365 AGCTTTACTGTGACGAGCGGTATGTTGTTTATTAAGCAACGCTCTTACGACTTGTGTG 424
419 ctgcgcgtcgt 478
425 CTACACGGGATATAGCTCCACGCGACATTTATATTAATCTGCAACTTTACTGCACTTGTG 484
479 cggcgagcccggt 538
485 CTGCGGATTTGGATTTCATTCGTTGAGGAGAAACGCGCCTCGAGACTGCATGATG 544
539 tattgagcgttggcgaagcgaagtcataatcagtggtgtgtgtgtgtgtgtgtgtgtgt 558
549 CCTTGGCAAACTCAGCAGGAAAGTGTATGTGGCTGGAAGCCGGGTTCATCTGTG 604
559 cgtcgaatctgtcgaacatcgaacatgaatctgcggcgccgctgcggcggtt 658
605 GCTTGCACAAATGCAAAATTCGTAGAAGTACA-----CGACCTCATCCGCGAGGTA 655

659 tgaatgacgacacattcattcatttgagcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 718
656 ACGTGGCGGTCTGTATCAATATAGAAAGCAATCAATAGCGGCGCAAGCGTGTGACGC 715
719 tcaattatcaagatgtaattccatcagcagcttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 778
716 TCAAGGCTACGCACTGATTCGATCGGACGCTTCTCTGTTACGGGCAAGCGGATTTTA 775
779 agcggtatgt 838
776 CCAGATGATTCGATGACGCGGCTCAGACGCTGACGCTCAGGATGATGCTCCGATTATGC 835
839 tgggt 892
836 CCGGTGCAATGCTTGTGCTTCTCTCCCGCGGACGCTGACATTAAGGATATCAGAGAC 895
893 gctatattccggt 952
896 GTGTGATGATGAGCAATGCTGACCGGTAAGAACTCTGCGAGAAAGGACCTTCTCTGT 955
953 gacgtaccacaatgaatattccgtatcgaagaagcgcg---agcaagagctgtgtg 1009
956 CAGCCCGGTGTACCAAGATCACGCTGATCCCGAAGCGCAGATGTGAGCAACTTTCG 1015
1010 gctgt 1069
1016 GGTGGCTCACCCCGCTCTCAATCTACAGCATGAGCAAGAGTATTTCTTGTGTTGC 1075
1070 ---tgaaaaacaactctcaagttcaagcagcgtcaacggtgtgtgtgtgtgtgtgtgtgt 1126
1076 AGGGGAAAAACAAGATGCTGATCTGATCCCGCATCAAGGCTGGGAGAGTGTGATGA 1135
1127 tgcgcatgt 1186
1136 TCATGACCAACAGATGACCGGCTTTCCTCCGATGACATCTATCCGAGATTTGTGTC 1195
1187 gcaattatcgt 1246
1196 AGGCTTTATGACATTCGACATGACATGACAAATGAGAGCACTTAAGCATATTAATGCTC 1255
1247 aagaagactcgt 1306
1256 CCGAGACATTTGCACTTGGAAATTTGTGATACATCAATCAATGACCTGCAGCTGATCG 1315
1307 tgcgtgaagtggt 1338
1316 TTTGCGATGCGCTTCAATGCTTCTATTAAGGA 1344

RESULT 18
HIN20229.1 4926 bp DNA linear BCT 09-FEB-1995
DEFINITION Haemophilus influenzae BOLA (bOLA), glutathione reductase (gor),
phosphatidylserine decarboxylase (psd), 30K protein (rpmf), genes,
complete cds.
ACCESSION U20229.1 GI:644850
VERSION 020229
KEYWORDS Haemophilus influenzae.
SOURCE Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
ORGANISM Haemophilus.
REFERENCE 1 (bases 1 to 4926)
AUTHORS Barack, G.J. and Helmer, S. R.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4926)
AUTHORS Barack, G.J.
TITLE Direct Submission
JOURNAL Submitted (24-JAN-1995) Gerard J. Barack, Biochemistry, University
of Maryland at Baltimore, 108 North Greene Street, Baltimore, MD
21201, USA
FEATURES
SOURCE 1. 4926
Location/Qualifiers

/translators="MTKHYIYIAIGGSGSGLASLRNRAAYGNKAALIEAKHGGTCVNV
 VGVCPKKMYFGAHIAEAINNNVAIDVGDEJVEAKPKDSKLIENSOAIIETSIHSYNVV
 LAKNRNDIVNGPCKRPVDNAHTLEVTLLADCTEDQUTADHILLATAGRPTRPRINIKOGEXIL
 DSVGFALLERLKPRAAYIGAGITLVNELSVNLNSLGERTHLLVRHRARQQDLILEVL
 LVEFVALQDGIQLTHNTSPSSIVKNADSDSLVRCGOSDVVDVCIMAGRVPYTDKG
 LEAGAVETHNGHYVKDYKONTNVKCIYAAGDIILENIELTPVAAGAARRLSERLPFNKN
 KPEFYEDSLVPTVYSHRPIGTGVGLEPALEIOYGAEINKVYSSTFTAMTATVTOHRR
 OPCRKLVCVGKDEKVYGLHGIGFVDEMIIOGFAYAIKMATKADFDNTVAIHPTGSESE
 EFTVMR"
 complement(3218. .4090)
 /gene="psg"
 complement(3218. .4090)
 /gene="psd"
 /function-"lipid metabolism"
 /codon_start=1
 /transl_table=1
 /product-"phosphatidylserine decarboxylase"
 /protein_id="AA62138.1"
 /db_xref="GI:644855"
 /translation="MNTSNYSWQRKYAFQVYMPQIYLTQIAGFAKOKMKTIHFVITY
 KATRKINIDMSIAQKEQSDYASFNEFFIRPLEKRNRPIONRTALCLPDGXISEC
 GHIDNNLLDAKHFFSLIEDLAEDKELVERFKGESVYTLLSPRDYHRAMPDDATL
 RKNITVYGDLESVNPFLAQHPVPLNFANERNRYCVDFEFQMWOILGAATTTAII GTTT
 MAGINIPRHNEVTKTWYEEBSA VKLKGOEMFMFOGSTVINLFQANOVRALDHLSVS
 NEYPRMGEILAYKK"
 complement(3364. .3374)
 /gene="psgd"
 /note="encodes DNA uptake recognition sequence"
 complement(33780. .33790)
 /gene="psd"
 /note="encodes DNA uptake recognition sequence"
 complement(3818. .3825)
 /note="encodes DNA uptake recognition sequence"
 complement(4097. .4100)
 /note="psgd"
 complement(4136. .4152)
 /note="OxYR recognition sequence (Toledano, M.B. et al.,
 Cell 78:897-909, 1994)"
 /bound_molecule-"Oxyr"
 4247. .4771
 /gene="rpmf"
 4247. .4771
 /gene="rpmf"
 /codon_start=1
 /transl_table=1
 /product="30k protein"
 /protein_id="AA62139.1"
 /db_xref="GI:644856"
 /translation="MQRVKLPYLVDPIKDAQRLDYGYUAANOLERLAESVNYVLSL
 AQVTLISFYDPQKLVMKGVKDVIDECOCENPYYKQTLCEFTYSRVAMDADDIDL
 PEIERPEIFENFGEIDILIGVEDELIALPLVPMHSBHCVSNOEGYVLPBELAKM
 KPNEFAVLANKOK"
 complement(4561. .>4926)
 /note="orf121"
 /codon_start=1
 /transl_table=1
 /product="unknown"
 /protein_id="AA62140.1"
 /db_xref="GI:644857"
 /translation="LOLRGVSEFHRIILYLIIQOOLKHYNDAYHDVNEIFCFVEQ
 IALLNLIFCFRIKLANANGRGFRANSNGNPKTSCSDTISQCSEDCICTGSKKAKISS
 SSTYKISGISPNASNGS"

	BASE COUNT	1598 a	862 c	938 g	1528 t
ORIGIN					
Query Match		7.5%	Score 101.4;	DB 1;	Length 4926;
Best Local Similarity		69.6%;	Pred. No. 5e-14;		
Matches 149;	Conservative	0;	Mismatches 64;	Indels 1;	Gaps 1

1 atgatataaatcaaaaggtcttaaacctgccaccgcggcagacccggagcaagtcatc 60
 ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 ATGATTACAAATTAAAGAGTTTGATCTGCCCAATTCGCAGAAACAGCACAGATTAAC 165

QY 61 tatgagcgccgcatcattaccgaagtcggttcgttcggaagaatccggtatgcg 120
 Db 164 CATACCGCAACGCTGTGATCAATGATGGATCTAGAGAGAGATGTGGGATGCGT 105
 QY 121 cccctgataagaatgaagaagcgatgcgcataaaagccgaagtctgttgaagc 180
 Db 104 CCTTCAATGAAGCTCGCGCAAGCGATGTTGAAAAAGGTCAGTACTTCTTGAAGC 45
 QY 181 aaaaagatccggcggtgtgttaccgcgcgc 214
 Db 44 AATCCGCTGCTAATTTTACAGCCCTG 12

AF429315/C 125020 bp DNA linear PRI 18-JAN-2002
 LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
 DEFINITION AF429315
 ACCESSION AF429315.1 GI:17646244
 VERSION
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 125020)
 HOLMES,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
 Ingber,S.L., Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
 Potter,N.T., Ross,C.A. and Margolis,R.L.
 A repeat expansion in the gene encoding junctophilin-3 is
 associated with Huntington disease-like 2
 Nat. Genet. 29 (4), 377-378 (2001)
 PUBMED 11694876
 MEDLINE 21583737
 JOURNAL 2 (bases 1 to 125020)
 REFERENCE
 AUTHORS Holmes,S.E., Ingber,S.L., Ashworth,R.G., Ross,C.A. and Margolis,R.L.
 DIRECT SUBMISSION
 TITLE Submitted (05-Oct-2001) Psychiatry, Johns Hopkins Medical
 INSTITUTIONS, 600 N. Wolfe St., Baltimore, MD 21287, USA
 LOCATION/Qualifiers
 source 1. 125020

FEATURES
 source 1. 125020
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="16"
 /map="16q24.3; between D16S520 and WI-12410"
 /note="Isolated from a patient with Huntington's
 disease-like 2 (HDL2)"
 repeat_region complement(35581..35746)
 /rpt_type=tandem
 /rpt_unit=ctg
 mRNA complement(35507..35687)
 /gene="JPH3"
 /product="Junctophilin 3"
 complement(35507..35687)
 /gene="JPH3"
 /note="JPH3"
 complement(35507..35687)
 /gene="JPH3"
 /note="Component of the junctional complex between plasma
 membrane and endoplasmic reticulum"
 cds complement(35507..35687)
 /product="Junctophilin 3"
 /protein_id="AA140941.1"
 /db_xref="GI:17646245"
 /translation="MSSGGRPNPDGSGYCGGWDKRAHGHGVCGRPKSGEYTGSM
 HPEVGLVTPWPGNTYGTWAGKRGHIGLESKRWYKGEWTHGFGKRYGVRECA
 NAKIEGTWNLDDGTETYSDG"
 BASE COUNT 29056 a 32731 c 30696 g 28283 t 4254 others
 ORIGIN

Query Match 3.4% Score 45.6; DB 9; Length 125020;
 Best Local Similarity 10.7%; Pred. No. 2.1;

Matches 97; Conservative 374; Mismatches 431; Indels 7; Gaps 2;
 QY 339 aagcgagcgaangantungnngcaatcgtatccatccggttgcgtgcgtan 398
 Db 17784 AGSKSRGCTGYMMKKGGGKSMSSKMKWGSSTSRSSAKSCSYMMGMSMCMSM 17725
 QY 399 ccgttcgttcgaataatccctgcgtgcgtgcgtgcgtgcgttcgttcgtc 458
 Db 17724 AKSYMMCYMYRMSSYMYKYCYSCMGMSSTSYSCCMKMSMSGCKMYTSGMS 17665
 QY 458 gatgacaccatccgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 518
 Db 17664 SKTSMGYTSTSTCKKYSMSMSMYCKTSKYRRKRSYYMMGCRRAKYYCAGRRR 17605
 QY 519 ttctgacgaangntgctggtatgagccgttcgtcgtcgtcgtcgtcgtcgtc 578
 Db 17604 -MSYMMCCAKMMWYSCWYCYMTYYKSCYKSSCYRRGGYMMWYKTSYSGSRSM 17546
 QY 579 ggcagctgcgcgaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 638
 Db 17545 YTCMRSSKSSKMSMSMASSMCWAGACGRRKRSAGMGARRSKGRSTGMRACSK 17486
 QY 639 cggccgcacatccggtcgttcgttcgttcgttcgttcgttcgttcgttcgtc 698
 Db 17485 KTGSGSTGRSMKKKGYSKYSGMKKKTKCYCMKKYKRTSMCMWYMMKSGYKRY 17426
 QY 699 aaacaaacggttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgt 758
 Db 17425 KRCMKKCCCTGTRGMSKSGYKSMRGMGSSYTSYSCWMSGYSYMMCKMKTSTKKR 17366
 QY 759 aacagcgcttcgaacacgcgcgttcgttcgttcgttcgttcgttcgttcgttc 818
 Db 17365 RSMRGSMSKMGWYAGRCYSBSMSSTRFRBSKYKSGRGMKMGWGMKRSKY 17306
 QY 819 acgcttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgt 878
 Db 17305 WSMKMKRSMKSKYSTSYSGRRKSKMGWSTSKSKASSKRMASKCTTSSYWSNN 17246
 QY 879 tgacgagacacgcgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgt 932
 Db 17245 RRNRMGKTGCMYRRSRAMMNGNAAAGCTCCCCANTNGGGGAAAAAGGCGSASR 17186
 QY 933 aagcgagcagatatttgcgttcgttcgttcgttcgttcgttcgttcgttcgttc 992
 Db 17185 ASCTKGMSKSCYRSGTRKMKSKCRGSRGSKSMGTRGSGGKTSYAKSGRCY 17126
 QY 993 cagcagaagcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 1052
 Db 17125 CWCWGRKGRKCMSSRMKMKYKYYRRKRMTCMKMYSMTYAMCRSMCMCKSCCGC 17066
 QY 1053 gacccctgcgcattcctgtaaaacacatcctcaagttcaagcgcgtcaaggtg 1112
 Db 17065 YSMGMSYSTSKYSWGMKYSYMSYKSTYSKSTYSKAMSSKRGMTGGRYGGGNSYCKG 17006
 QY 1113 cgaagcgcatgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 1172
 Db 17005 GSYKGGGSMGKGGKSRGSSAGKSSYKMKSCARYSKYSCYSRMMCMMSYTCWG 16946
 QY 1173 taccgcttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 1232
 Db 16945 GKGMSGMSMRKSSGSGWAGWMDGSYTDKSDKMSRBSBYKSKWASRRGKY 16886
 QY 1233 ctggaatt 1241
 Db 16886 YTCDSMKK 16877
 RESULT 21
 LOCUS BACGNTZA 36200 bp DNA linear BCT 05-FEB-1999
 DEFINITION Bacillus subtilis 36kb sequence between gnt2 and trnY genes
 ACCESSION D78193

Query Match	Best Local Similarity	Score	DB 1:	Length	DB 2:
1	100.0%	100.0	DB 1:	4926	4926
224	99.9%	99.9	DB 1:	4926	4926
Matches	153	Conservative	0	Mismatches	69
			Indels	1	Gaps
					1

3.58; Score 47.4; DB 9; Length 125020;
11.4%; Pred. No. 0.76;

	Matches	45	Conservative	172	Mismatches	177	Indels	0	Gaps	0
QY	239	accgcgggaaagagcgctactcaatgcagtcgcatgcatgctcgttgaaagcaacgaa	298							
Db	17040	ksfmrskmcsmrsmcsrsrskcsrccggsmgkcgkcgksgtgrtkrkskgnkkgakwm	17099							
QY	299	tccagattgaagctagctaccctgaagcgctgycgaactaaagcgcgaagaagtcgc	358							
Db	17108	yyrrsmnrkmmkmtyskgmcmcmcgrrgcycscmtsrsmcscycakcksmcysfvg	17159							
QY	359	gcaacctgatacaatccggttgctggaactgcgcgcgcaccgcctcgcgtacgaacattc	418							
Db	17160	msnkgtytactcyrsgsmsskcmrsgyststscgcccctttttccccnntggcaacttt	17219							
QY	419	ctgcgcgtcagatcgagcgcgttcgccatcttcgtcaatgcaatgagacaaatcgcgtg	478							
Db	17220	tnmknkyyeyrrkngcamckyrnnynnmwmsrscragmsctrkysntmtmsaycwmcmw	17279							
QY	479	cgcgcgaacctcagctcattcaaaagaagcgcgcggattttaaagcgcgcgttg	538							
Db	17280	ytcsmsnmasrcmmsstymkmskswrnsctymkcmcmkncfcmrsmrsgmsytmfas	17339							
QY	539	tattgagcgttttgacgcgaacgaatccatcgtttgttgaagcagctgtgcgacagctgc	598							
Db	17340	wkssscgctctcrtkcmcmsskscyrstymmskrmgcmkmsrcwmgsnmgasrbscrtcy	17399							
QY	599	cgcttgaaatctgctgcacatcgcgaacacatga	632							
Db	17400	ksmrcsmsskcyrcaccmkmkgymrmrcsm	17433							
RESULT 21										
BACGNTZA/C										
LOCUS										
DEFINITION										
ACCESSION										
VERSION										
KEYWORDS										
SOURCE										
ORGANISM										
REFERENCE										
AUTHORS										
TITLE										
JOURNAL										
REFERENCE										
AUTHORS										
TITLE										
JOURNAL										
MEDLINE										
REFERENCE										
AUTHORS										
TITLE										
JOURNAL										
MEDLINE										
REFERENCE										
AUTHORS										
TITLE										
JOURNAL										
MEDLINE										
REFERENCE										
AUTHORS										
TITLE										
JOURNAL										
MEDLINE										
REFERENCE										
AUTHORS										
TITLE										
JOURNAL										
MEDLINE										
REFERENCE										
AUTHORS										
TITLE										
JOURNAL										
MEDLINE										
REFERENCE										
AUTHORS										
TITLE										
JOURNAL										
MEDLINE										
REFERENCE										
AUTHORS										
TITLE										
JOURNAL				</						

Victoria 3052, Australia

Location/Qualifiers	Location/Qualifiers
1. .1356	1. .1356

aps

```

/organism="Porphyromonas gingivalis"
/strain="W50"
/db_xref="taxon:837"
1..1356
/feature_start=1

```

```

/trans1_cable-11
/product="50 kD antigen PGI"
/protein_id="AAD33930.1"
/db_xref="GI:4929280"

```

BASE COUNT	338 a	333 c	380 g	305 t
ORIGIN				

Query Match	94%;	Score 126.4;	DB 1;	Length 1356;
Best Local Similarity	46.78;	Pred. No. 4,4e-21;		
Matches 631;	Conservative	0;	Mismatches 691;	Indels 30;
				Gaps 6

Db	11	TTATAAAAAA	AAAAAGGCGCTTGCATCTTAATCTGAAGGAATACCGCGCCCGAGATCC	70
Qy	62	atgacgccccggccc---	attaccgaagccggcttgcttgaggaagaatatgtccgacatgc	118
Db	71	TGGCGGAACCGGCCCCAAGTCTTACTTACGCGCTGTGCTGCCACGATTTTGAAGGTGA	130	
Qy	119	gccctcgatgaaataccaaggaaggtgaagccgctggaaaaaagccaagtgtcttgaag	178	
Db	131	TCCCCAAAGTGACGGCTGTGTCGGGGGATTAAGTGTGCTCCGGCTTCAGCATGATCCACC	190	
Qy	179	acaaaagaatccggcgtagatattactgcccggctccaggaacaaatccggccatc	238	
Db	191	ACAAAGCATATCCGAGATGAAGTTTAAAGTCCGGTTAGCGCGGAAGTATCCGGGTGA	250	
Qy	239	accgtgcgaaagcgctacttcgttcgctgcgtgattgcgtttgaagcaacgacgaa	298	
Db	251	ATCGGGTGCCAAAGCGCAAGGTGTGAGCATGAGAGGTGAACCGAOCGGACTGACGAAT	310	
Qy	299	tcgaattcgaaagcctacgttacctgaagcgcgtgcaaatgagcaggaagaagtgcgc	358	
Db	311	ACGAATCATTTCCCTGTGCGGGAT-----CCGTCTGCCCTCTCTGCCGAAACAGATCAAG	364	
Qy	359	gcaacctgatcaatcaatgaagctatgactgactgcgcaccgctccgttcagaataatcc	418	
Db	365	AGCTTTATCTGCAAGCGGTATGTGGGTTTATTAAACAACCTCTTTAGACATTAATGG	424	
Qy	419	ctgcgcgtaatgcccgcgcgcttcgcacatcttcgtaatgtagaagaaacaacatccgttgc	478	
Db	425	CTAACCGGATATAGCTCCACGCGAGCAATTATTATTTACTGCCAACTTATCGCAACGATTCG	484	
Qy	479	ctgcgcgcacctacgactcatcaaaagaacgcgcgaagaacttcaaaacgcgcctgttgc	538	
Db	485	CTCCGGAGCTTCGATTTCATTCGTTCGAGGAGAAAGAACCGCCCTCCAAACATGCCATCGATG	544	
Qy	539	tatttggccgcctgacgcgaagcttaaatccatctgtgttaagcaggaagcgccagacgtgc	598	
Db	548	CCTTGGCCAAATCTACGACAGGAAGGTGATGTGGGCTCGAACCGGTTCAATCTCTGG	604	
Qy	599	cgctcgaaaatgtctgccaatatcgaaacacatagaatttgcgcgcgcgcacatccctgcgct	658	
Db	605	GCTTGGACATTCGAGAAATGTAAGATTAACA-----CGAGCTTCATCCGGAGGTA	655	

QY 659 tggatgacagacacattccattcatctcgagccagctcgscggaataaacgcttggagca 718
 Db 656 ACGTGGCGCTGTGATGATTAATCATACGAAGCCATTCAATCGGGCGGAACGCTGTGAGCG 715
 QY 719 tcaatttcagaagcgtgatattcgtactatcgagacgcttctgtcaacagcgcgtctgaatccg 778
 Db 716 TCAAGGCTACCGACACTGATGCTGATCGAGATCGATGTTCCGCTTTACGGGGAAGCCGATTTTA 775
 QY 779 aagcgctgattgctcttggggcgcgctcgtaagtaacaacacgcgcctcttggtaacgctt 838
 Db 776 CCAGAAATGATGGCCCATACCGGCTCAGACCGCTGCAGCTCAAGATACGTCGGATTTATGC 835
 QY 839 tgggtgcgaagctgtctcacaattcacgcgcggcggaattgct----gacgcgagcaac 892
 Db 836 CGGCTGTGCAATGCTTTTGCTCTCTTCCCGGCGCAGTACACAAAGAAATCTCACAGC 895
 QY 893 gacgtatctccgcttgcgtatgtatgaagctgtgatctgaacaagscgcgcatgatattcg 952
 Db 896 GTGTGATCGATGGCAATGTGCTGACCGGTAAAGAGTCTCGAGAGAAGAGCTTTCTGT 955
 QY 953 gacgtaccacacatcagaattccgttalcagaagaagccgcgagca--aagagctgtcg 1009
 Db 956 CAGCCCGGTGTGACACAGATCAAGGAGATCCCGAAGCGAGCATGTGGACGAACCTCTTC 1015
 QY 1010 gctgggttgcgcgcgacgcgagacaagaatctccalcacagcgcacacactctgagccattcc 1069
 Db 1016 GGTGGCTGCGCACCCCGTCTCGATCGTACACACATGAGCAGAGCTTATTTCTTGGTTGC 1075
 QY 1070 ---taaaacaacactctcaagttcacaagcagacgcgtcaacgcgcgcgcgcgcacatg 1126
 Db 1076 AGGGAAAAACAAAGATGATGATCTACTCATGTCCCGGATCAAGAGGTGGCAAGCTGTATGA 1135
 QY 1127 taacgcgcgagcactatgagcgcgtaatgctgtgagacatccctactcctctgcttctgc 1186
 Db 1136 TCATATGCAACAGATGATACCCGCTGTTTCCGATGCAATCTATCTCCGAGATTTGTGCA 1195
 QY 1187 ggcatttaatcgctcgagatataccagcaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1246
 Db 1196 AGGCTATTATACATTCCGATCGATCGAAGATGAGAGACTTATGCAATATGAAAGTGGCTC 1255
 QY 1247 aagaagacgcctctgtgtcagctcgtctgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1306
 Db 1256 CGGAGAGCTTTCGCACTTCGAAATTTGTGATGATATCCAAAGATCAGAGTGCAGCTATG 1315
 QY 1307 tgcgcgaagtgcctggaacacatctggaagaagaa 1338
 Db 1316 TTGCGACACCGCTTGAATGCTCTATTAAGGA 1347

RESOLUTION 119
1990-03-28/0

LOCUS	4926 bp	DNA	linear	BCF 09-FEB-1995
DEFINITION	Haemophilus influenzae BOLA (bola), glutathione reductase (gor),			
	30x protein (rmpf) genes.			

phosphatidylserine decarboxylase (psd), von protein (4 pmol / gene)
complete cds.

ACCESSION	U20229
VERSION	U20229.1
	GI:644850

KEYWORDS :
SOURCE :
Haemophilus influenzae.

ORGANISM Haemophilus influenzae
bacteria: proteobacteria: gamma subdivision: Pasteurellaceae:

Haemophilus.

REFERENCE	AUTHORS
1 (bases 1 to 4320)	Barcak, G.J. and Helmer, S.R.

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4926)

AUTHORS Barcak, G. J.
TITLE Direct Submission

JOURNAL Submitted (24-JAN-1995) Gerard J. Barcak, Biochemistry, University of Maryland at Baltimore, 108 North Greene Street, Baltimore, MD

Ol Maryland at Baltimore/ 100 North Second Street
21201, USA
Telephone (410) 441-1100

FEATURES	LOCATION/Qualifiers
SOURCE	1. .4926


```
/organism="Haemophilus influenzae"
/strain="Rd"
/isolate="KW20"
/db_xref="taxon:727"
complement(27..224)
/note="Orf65"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AA62134.1"
/db_xref="GI:644851"
/translation="MITIRKGLDLPFAGKPAQVHSGNANVOAILGEIYVGRPSMK
VREDVYVKKGVLEDFEDKKILY"
complement(519..842)
/gene="hola"
complement(519..830)
/function="stationary phase regulatory protein"
/transl_table=11
/codon_start=1
/product="BOLA"
/protein_id="AA62135.1"
/db_xref="GI:644852"
/translation="MSIOGIIROKIOKEPPIAIEENSRLHNSRGSSEHKCVIY
SADRNIKRVQHRQRIYQLINELNHSIALMLHFTPEMKRQNETVPHSTKCAIG
R"
misc_feature
775..785
/note="encodes DNA uptake recognition sequence"
complement(838..842)
/gene="hola"
RBS
877..913
/note="OxyR recognition sequence (Tartaglia, L.A., et al.,
J. Mol Biol. 210: 709-710, 1989)"
/bound_molecy="OxyR"
915..919
/note="orf199"
924..1523
/note="orf199"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AA62136.1"
/db_xref="GI:644853"
/translation="MALVNRKIKTLSSVGLATLFLAGCOASNITLFTPPAPSAMN
VRIAVSVYTRDSRAIOELASTYKRGELIKLASPSYQLFOOVMOONLSKGFVYG
QINGSRMYTVDYRREGTGYEGCNLRKIKTKQAPVYVYGAKGSKNSNVTHSQSG
VFNAGDELHKVLSQTFNDIVNNITDDQEVAAAINOYSN"
1343..3109
/gene="gor"
1343..1348
/gene="gor"
/note="putative"
1365..1370
/gene="gor"
1365..1370
/gene="gor"
/note="putative"
1585..1590
/gene="gor"
1605..1610
/note="putative"
1605..1610
/gene="gor"
/note="putative"
1728..1731
/gene="gor"
1739..3109
/gene="gor"
/EC_number="1.6.4.2"
/function="reduces oxidized glutathione using NADPH"
/codon_start=1
/transl_table=11
/evidence="experimental"
/product="glutathione reductase"
/protein_id="AA62137.1"
/db_xref="GI:644854"
```

```
/translation="MKRHYDIAIGGSGGSIASLNRASTGKCAITIEAKHIGTCVN
VGCWPKVMVEYGAHIEAINNYPADYGFDEVEVKFEDSKLIESQATISRLHSYNNV
LAKNIDIVNGFGKPVDAHILEVTADGTEKVOYADHILATGGRPYRPNKIGQEGYI
DSDFPALTELPRAAVIGAGYLAVELSGVNSGCGEDVDCVYIAGAAVRNDPILVER
LYEVLAODGIDGHTNSTPSEIVKNADSLVRCGSGEDVDCVYIAGAAVRNDPILVER
LENAGVETNENHGVKDKYKNTYKGIYAGGDIENGLITLTPVAAGRRLSERLFNN
KPEYLDYSLVPTVFSHPITGYGLTEPAIEQYGAENKVTYSSSTFAMATYATQHR
QPCMKILVCVKDEKVEVGLHIGFVDEMTQGFVAVAIKMGATKADPNTVAIHPTGSE
EFVTMR"
complement(3218..4090)
/gene="psd"
complement(3218..4090)
/gene="psd"
/function="lipid metabolism"
/codon_start=1
/transl_table=11
/product="phosphatidylserine decarboxylase"
/protein_id="AA62138.1"
/db_xref="GI:644855"
/translation="MMTSNYSWORLKAFOYVMPQIYLTQIAGFAPKQKMGKITHFVI
KAPAKINIDMSIAKQFSDYASFNEFEIRPLEKEMRPIQNPALCLPADGRISEC
GHIDNILLQAKGFESLEDLADKEIVETFKNGEPYTYLSPDYRHRVMPDQATL
RKMIYVGDLEFVNPLAQHVPLFARNERYICVFDFGTWQILVATITTAAGT
WAGVYNPRHNEVKTWTYEGESAVALKLGQEMGFQLSGTVINLFOANQVRLADHLSV
NEPVAMEILAYKK"
complement(3364..3374)
/gene="psd"
/note="encodes DNA uptake recognition sequence"
complement(33780..33790)
/gene="psd"
/note="encodes DNA uptake recognition sequence"
3818..3825
/note="encodes DNA uptake recognition sequence"
complement(4097..4100)
/note="psd"
complement(4126..4152)
/note="OxyR recognition sequence (Toledano, M.B. et al.,
Cell 78:897-909)"
4247..4771
/bound_molecy="OxyR"
4247..4771
/gene="rpmf"
4247..4771
/gene="rpmf"
/codon_start=1
/transl_table=11
/product="30k protein"
/protein_id="AA62139.1"
/db_xref="GI:644856"
/translation="MOKYKPLPLVDPPIKNOBRIDVYGYANQLERLAESYVNLSD
AOVTLSTYDPOKLIVMKRQVIDDLECORCNEPKYKOTLCEFTYSPVANDQADL
PEIYPIEFENERGEIDLIGTVEDELIALPLVPMHSSECEVSAQOVGVLPBELAK
KENPFAVLANLNRK"
complement(4561..4926)
/note="orf121"
/transl_table=11
/codon_start=1
/product="unknown"
/protein_id="AA62140.1"
/db_xref="GI:644857"
/translation="IOLRGVSEFFRLYLILIOQULLKHRVNDAYHDVNEIFCFVEOP
LAILNLIVYCFRLANTANGFGFANSSGNTPTSCADPSCDSCDSCIGTSGRAKISS
SSTVPKTSIPNSINSIGS"
BASE COUNT 1598 a 862 c 938 g 1528 t
ORIGIN
Query Match 7.8%; Score 105.4; DB 1; Length 4926;
Best local Similarity 70.0%; Pred. No. 9,7e-16;
Matches 156; Conservative 0; Mismatches 66; Indels 1; Gaps 1;
```

```
QY 1 atgattaaatacaaaaagcttaaatctgccatcgcgagcgagcgagcaagtcatt 60
Db 224 ATGATTCAATTAAGAAAGGTTTGATCTCCCAATTGGGGAAAAACGACAGTAATTC 165
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: June 30, 2002, 12:42:41 ; Search time 27321.5 Seconds

(without alignments)
3375.485 Million cell updates/sec

Title: US-09-303-518d-653

Perfect score:

Sequence: 1 atgaacacacgacacacg.....attagctaccgtgtaa 4407

Scoring table:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

GenBank:
1: gb_ha.*
2: gb_hg.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_ph.*
7: gb_pi.*
8: gb_pi.*
9: gb_pi.*
10: gb_pi.*
11: gb_pi.*
12: gb_pi.*
13: gb_pi.*
14: gb_pi.*
15: gb_pi.*
16: gb_pi.*
17: gb_pi.*
18: gb_pi.*
19: gb_pi.*
20: gb_pi.*
21: gb_pi.*
22: gb_pi.*
23: gb_pi.*
24: gb_pi.*
25: gb_pi.*
26: gb_pi.*
27: gb_pi.*
28: gb_pi.*
29: gb_pi.*
30: gb_pi.*
31: gb_pi.*
32: gb_pi.*
33: gb_pi.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	ID	Description
--------	-------	-------	-------	--------	----	-------------

C	73	135.6	3.1	51334	1	AF200692
C	72	135.6	3.1	48254	6	AX276220
C	71	135.6	3.1	4394	1	AF097644
C	70	135.6	3.1	4113	6	AX276281
C	69	135.6	3.1	5096	1	HIGA
C	68	140	3.2	5096	1	HIGA
C	67	140	3.2	21494	1	SEPRIN
C	66	141.6	3.4	5719	1	AF369907
C	65	152.8	3.5	5088	1	HEAIGALPTB
C	64	152.8	3.5	15212	1	U32779
C	63	152.8	3.5	5088	1	HEAIGALPTB
C	62	152.8	3.5	5088	1	U32779
C	61	152.8	3.5	5088	1	HEAIGALPTB
C	60	152.8	3.5	5088	1	U32779
C	59	152.8	3.5	5088	1	HEAIGALPTB
C	58	152.8	3.5	5088	1	U32779
C	57	152.8	3.5	5088	1	HEAIGALPTB
C	56	152.8	3.5	5088	1	U32779
C	55	152.8	3.5	5088	1	HEAIGALPTB
C	54	152.8	3.5	5088	1	U32779
C	53	152.8	3.5	5088	1	HEAIGALPTB
C	52	152.8	3.5	5088	1	U32779
C	51	152.8	3.5	5088	1	HEAIGALPTB
C	50	152.8	3.5	5088	1	U32779
C	49	152.8	3.5	5088	1	HEAIGALPTB
C	48	152.8	3.5	5088	1	U32779
C	47	152.8	3.5	5088	1	HEAIGALPTB
C	46	152.8	3.5	5088	1	U32779
C	45	152.8	3.5	5088	1	HEAIGALPTB
C	44	152.8	3.5	5088	1	U32779
C	43	152.8	3.5	5088	1	HEAIGALPTB
C	42	152.8	3.5	5088	1	U32779
C	41	152.8	3.5	5088	1	HEAIGALPTB
C	40	152.8	3.5	5088	1	U32779
C	39	152.8	3.5	5088	1	HEAIGALPTB
C	38	152.8	3.5	5088	1	U32779
C	37	152.8	3.5	5088	1	HEAIGALPTB
C	36	152.8	3.5	5088	1	U32779
C	35	152.8	3.5	5088	1	HEAIGALPTB
C	34	152.8	3.5	5088	1	U32779
C	33	152.8	3.5	5088	1	HEAIGALPTB
C	32	152.8	3.5	5088	1	U32779
C	31	152.8	3.5	5088	1	HEAIGALPTB
C	30	152.8	3.5	5088	1	U32779
C	29	152.8	3.5	5088	1	HEAIGALPTB
C	28	152.8	3.5	5088	1	U32779
C	27	152.8	3.5	5088	1	HEAIGALPTB
C	26	152.8	3.5	5088	1	U32779
C	25	152.8	3.5	5088	1	HEAIGALPTB
C	24	152.8	3.5	5088	1	U32779
C	23	152.8	3.5	5088	1	HEAIGALPTB
C	22	152.8	3.5	5088	1	U32779
C	21	152.8	3.5	5088	1	HEAIGALPTB
C	20	152.8	3.5	5088	1	U32779
C	19	152.8	3.5	5088	1	HEAIGALPTB
C	18	152.8	3.5	5088	1	U32779
C	17	152.8	3.5	5088	1	HEAIGALPTB
C	16	152.8	3.5	5088	1	U32779
C	15	152.8	3.5	5088	1	HEAIGALPTB
C	14	152.8	3.5	5088	1	U32779
C	13	152.8	3.5	5088	1	HEAIGALPTB
C	12	152.8	3.5	5088	1	U32779
C	11	152.8	3.5	5088	1	HEAIGALPTB
C	10	152.8	3.5	5088	1	U32779
C	9	152.8	3.5	5088	1	HEAIGALPTB
C	8	152.8	3.5	5088	1	U32779
C	7	152.8	3.5	5088	1	HEAIGALPTB
C	6	152.8	3.5	5088	1	U32779
C	5	152.8	3.5	5088	1	HEAIGALPTB
C	4	152.8	3.5	5088	1	U32779
C	3	152.8	3.5	5088	1	HEAIGALPTB
C	2	152.8	3.5	5088	1	U32779
C	1	152.8	3.5	5088	1	HEAIGALPTB

Db 1261 GAGGTAATTTTACGGTCTCCGCTTAAACAAAGAAAGCGGCAAGCGCGGCGCTTCAT 1320
QY atcagtgatgagagatccgcttaacttgaaagtaaacgagctgcaaacgacgcctgtcc 1380
Db 1321 ATCAGTATGAGTACCGTACTTACTTGAAGTAACGGGCTGGCAACGACGCGCTGTCC 1380
QY 1381 aaaaatgcaaaagcagcgtctgctgcttcaagcgaaggaagaaacagagctgctgagc 1440
Db 1381 AAATTCGCGCAAGGCAAGCGCTGCTGCTTCAAGCCAAAGGGAACCAAGGCTCGTCCAC 1440
QY 1441 gtaggagagagatgcatcttagatgagcagcgagcagatcaaggaacaaacagcc 1500
Db 1441 GTGGGCGCAAGGTAATCTTATGATGAGCGAGCGGAGATCAAGGCAAAACCAACCC 1500
QY 1501 tttagtgaatccgctgtgctgagcagcgagggagcgtagcaactgactgacgataatcag 1560
Db 1501 TTTAGTGAATCGCTTGGTTCACGCGGAGCGGAGCGCTGCAACTGATTCGATATCAG 1560
QY 1561 ttaaccccgcaaaactctattcagcttccgagcgagccttgatgattgaacgagat 1620
Db 1561 TTCAACCCCGCAAACTCTATTTCGCTTTCGCGGCGAGCTTGGATTGCAAGCGGCTAT 1620
QY 1621 tgcgcttcgctcagcagcagatcaaaatcacgagtgaaagggagagatgactgcaacacat 1680
Db 1621 TCCGTTTCGTTCCACCGCATTTCAAAATACGATGAAGGGGCGATGATGTCACCCCAAT 1680
QY 1681 caagacaagaatccacgcttaccatcagcagcaataaagatattaccacaacgagcaat 1740
Db 1681 CAAGACAAAGAAATCCACCGTTACCATTTACAGGCAATMAAGATATTACTACACCGGCAAT 1740
QY 1741 aacaaacttgatgagcaaaagaatgtgctcaaaagctggttggggagagaagat 1800
Db 1741 AACAACTTGTGATGCAAAAGAAATGCTTACACAGGATGCTTGGGCGAGAAAT 1800
QY 1801 gcaacaaagaagcagcgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1860
Db 1801 GCAACCAAAAGCAAGCGCGCTCAATCTGAATTACCAACCGGAGAGCGGATGCGACT 1860
QY 1861 ttaactgcttcgagcagcaaaatctaaacgagcaataacagcaaaacgagcaactg 1920
Db 1861 TTACTGCTTCCGCGGAAATTTAAACGGCAATATTCACGCAAAACCGGCAATCTG 1920
QY 1921 ttttaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1980
Db 1921 TTTTTCAGCGGAGACCGGACCGGACCGCTTCAATCATTTAGAAAGCGGCTGTCAAAA 1980
QY 1981 atggaagatctcccaagagagaaatcgctgtggagacagcagcagcagcagcagcagc 2040
Db 1981 ATGGAAGGTATCCCAAGAGAAATCGTGGGACAAAGATTTGATGACCGGCAATTT 2040
QY 2041 aaagcgaagaaactccatctcagcagcagcagcagcagcagcagcagcagcagcagcagc 2100
Db 2041 AAACCGGAAACTTCCATATTCAGGCGGCAAGCGGCTGTTCCCGCAATTTGCCAAA 2100
QY 2101 gtaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2160
Db 2101 GTGGAAGCGATGCGATTTAAAGCAATCAGCGGAGCGGCTTGGCTGTCCGACGCGAT 2160
QY 2161 caaagcacaacatctgtacagcagcagcagcagcagcagcagcagcagcagcagcagc 2220
Db 2161 CAAAGCCCAACAACTGTACAGCTTGGAGTGAAGCGGCTGTAACAAGTTGTAACCAAAA 2220
QY 2221 accatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2280
Db 2221 ACCATTTACCGACGATTAAGATGCTTCAATGACCAACACCGGATCAGAGCGCAATGTC 2280
QY 2281 agccttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2340
Db 2281 AGCCTTGGCGATCAGCTCTATTTAAATCTCAGAGGAGCTTGCCACACTCAACGCGCATCT 2340
QY 2341 agtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2400
Db 2341 AGTCAGGCGAGACGCGACTATACGGTTACGCGCACCGCCCAAAACGCGCAACCTC 2400

QY 2401 agcctcgtggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2460
Db 2401 AGCCTGTCGGGCAAGTCCCAAGCAACATTTAATCAACCAATTTAAACGCGCAACATCTC 2460
QY 2461 gcttcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2520
Db 2461 GCTTCGCAACATGCTTCAATTTATCTAAGCAACAGCGGCTGACAAACGCGGCTGTACG 2520
QY 2521 ctttcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2580
Db 2521 CTTCGCAACGCTTCAAGCAACGCTAAGCAATTCGACTCAACGCGGCAATGCTCTCA 2580
QY 2581 gccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2640
Db 2581 GCCGATTAAGCGATTTATTCATTTGAAACAGCGGCTTACCGGCAAAACGCGGCGG 2640
QY 2641 aaggaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2700
Db 2641 AAGGATACCGCATTTACCTTAAAGCAACGCAATGAGACGCGGCTGCGGCGGCAATTA 2700
QY 2701 ggcgaattaaaccttgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2760
Db 2701 GGCATTTAACTTGAACGCAAGCGCACTTACCTCAATTCGCTTATCAGACAGATGCG 2760
QY 2761 gcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2820
Db 2761 GCAGGCGGCAAAACCGGCGAGTGCAGCATGCGCGCGCGCGCTTGCCTTCCCTA 2820
QY 2821 ttaactgcttcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2880
Db 2821 TTACTGCTTACCGCGCAACTTCCGCAAGATGCGGCTTCAACAGCGGATGATTAACGCG 2880
QY 2881 aaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2940
Db 2881 AAATTTAAGCGTACAGGAACTTCCGCTTATGTCGGAATCTTTCGCTTACCGGAGGCG 2940
QY 2941 aaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3000
Db 2941 AAATTTAAGCGGCAAGGTTCCCAAGGCACTTACCTTGCCTTCAACATATCCGCG 3000
QY 3001 aacgaacccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3060
Db 3001 AACGAACCGGTAAGTCTGAGCAATTTGAGGATGAGGAAAGAAAGCAACACGCGCTG 3060
QY 3061 tccgaatcttaacttaacttaacttaacttaacttaacttaacttaacttaacttaact 3120
Db 3061 TCCGAATCTTATTTATTTACCTGCAAAACGACAGTGTATGCGGCTGTAT 3120
QY 3121 cagctatccgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3180
Db 3121 CAGCTTATCCGCAAGACGCGGAGTTCGCTGCTTAATTCGCTTCAAAAGAAAGAGACTT 3180
QY 3181 tccgaacacacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3240
Db 3181 TCCGAACAACTGCGGCAAGCGGAGGAAACAGAGCGCGCTTGAAGGCAACAGGCA 3240
QY 3241 ctggcgcgaac 3300
Db 3241 CTGGCGCAACACAGCGGAGGAAAGCAAGCGGCTTGAAGCGGCTGTAT 3300
QY 3301 gccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3360
Db 3301 GCCGCGGCAAGTCCCAAGGCAAGGAGGAAAGTGTGCGGCAAGCGGCGGCAAGCGG 3360
QY 3361 ggggaataatccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3420
Db 3361 GGGGAATAATCCGCGCTTATTCAGCGGAGGAAAGGAAAGGCGGCTGAGCGGATATA 3420
QY 3421 gacacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3480
Db 3421 GACACGCGCTTGGCGAAGCAAGCGGCAAGCGGAAACCGCGGCTTACCAACCGCTTCC 3480

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS
RESULT 2				
A96616	Sequence 649 from Patent WO9924578.	A96616.1	GI:6780200	unidentified.
LOCUS	4374 bp	DNA	linear	PAT 07-SDP-2000
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
unidentified.				

[illegible]

```
Db 781 GAGGCTCATTGGCGACAGTGCTCACCAATGTTATCTATGATGCCCAAAACCAAAAG 840
Oy 841 Tggttaataatgaggtatctgaacaggaacccctatacagaagaaagatggtctc 900
Db 841 TGGTTAATTAATGGGTATTTGCCAAAGCGGCAACCCCTATATAGAAAGCAATGGGCTTC 900
Oy 901 caagtagctcgtaaagaatggtctatgataagaaatcttgctggagagataccatcagta 960
Db 901 CACTGCTGTTGTAAGATTGGTTCTATGATGAATCTTGCTGGAGATACCATTCCAGTA 960
Oy 961 tctacgaacacacatacaatactctctacactatacagtaaaacagacgct 1020
Db 961 TTTCTAGAACACGCTCAAAATGGGAATTCCTTTTAAACGAGATAAATGAAGGCAACAGA 1020
Oy 1021 aaaaatgacgacacatacaatactctctacactatacagtaaaacagacgct 1080
Db 1021 AAATATCATGCGCAAAACATGACACATCTCTGCTTAATAGATTAATAACAGAAACGCTT 1080
Oy 1081 caattgttaatggtctctctacacagacagaagaagaacggttatacgtcaggt 1140
Db 1081 CAATTGTTAATGTTCTTTATCCAGACAGCAAGAACTGTTTATCATGCTGCAAGT 1140
Oy 1141 ggggtcaacagatatacagccagactgaataatgagaataatctcttatacgaacaa 1200
Db 1141 GGGTCAACAGGTATGACCCAGACATGAATATGAATAATTTCTTATATGACGAA 1200
Oy 1201 ggaagaagtgaatltgatactacacgaacacacagacgagcggtgttattt 1260
Db 1201 GGAAGAAGGCAATGATGACTTACACCAATCATCAACAGTGTGAGAGATATATATTC 1260
Oy 1261 gagggttaattttagctctcgcctaaacagaagaagctggcagaagcgaggtctat 1320
Db 1261 CAAAGGAGATTTTACGCTTCGCTGCTGAAATTAACGAACCTTGCAAGGCGCGGCTTAT 1320
Oy 1321 atcaatgatacgtacgtctacacttggaagtaaacgagcggtgacacagacgctctcc 1380
Db 1321 ATCAGTAAAGACAGTACCGCTTACTTGAAAGTAACGGGCTGGCAACGACGCGCTGACC 1380
Oy 1381 aaaaatgacgacacacacgctctgctcgaagcagaaggggaaacacagacgctcgtcgc 1440
Db 1381 AAATCGCGCAAAAGCAACGCGCTGCTTCAAGCAAAAGGGGAAACCAAGGCTGATCAGC 1440
Oy 1441 gtaggagacggttaagtaactcttagatcagagagcgagacatcaaggggaaaaaagacg 1500
Db 1441 GTGGGCAACGCTACAGTCAATTTGGATCAGACGACGACGATAAAGCAAAACAAAGCC 1500
Oy 1501 tttagtgaaatcggtctgctcagcgagcgagggagcggtgcaactgaatgacgatacag 1560
Db 1501 TTTAGTGAATCGGCTTGGTCAAGCGGAGGCTACGCTCAACTGAATGCCATATATCAG 1560
Oy 1561 ttcacacccgacacacactctatcttcggtctctcgcgcgagacgcttggaattgagcgagat 1620
Db 1561 TTTCAACCCCGCAAAACGCTTTCGCGCTTCCGGCGGAGCTTTGGATTTAAACGCGGAT 1620
Oy 1621 ttcgcttcgctcgaacgacatcaaaaataccgataagagggcgatgatactgtaacacacat 1680
Db 1621 TCGCTTTCTGTTCCACCGCTATTTCAAAATACCGATGAAGGGCGCATGTTCTCAACACAT 1680
Oy 1681 caagacagaagaatccacgcttaccatacagacataaagataatacactacaacgagacat 1740
Db 1681 CAAGACAAAGAAATCCACCGCTTACCATTTACAGGAATGAATTTGCTTCAACACCGGCAAT 1740
Oy 1741 aacacacacttgatagcaaaaagaagaatctgcttacaacggtctggtcttgagagaagaat 1800
Db 1741 AACACACGCTTGGATAGCAAAAAAATGCTCTACACAGTTGGTTGGCGAGAAAGAT 1800
Oy 1801 gcaacccaacgaacgagcggtcgaatctgaatctaccacacgagaagaacggaatcgaact 1860
Db 1801 ACGACCAAAAGCAAGCGGCGCTCAACCTTTTACACACCGCGCCAGAAACGCGACAC 1860
Oy 1861 ttaactgcttcgcgggagacaaatcttaacggaataatacagcaaaacgaagcgaactg 1920
Db 1861 ttaactgcttcgcgggagacaaatcttaacggaataatacagcaaaacgaagcgaactg 1920

Db 1861 CTGCTGCTTTCCGGCGGACAAATTTAAACGGCAACATCAGCAACAAACGGCAACCTG 1920
Oy 1921 tcttcacgagcagaacacagacacgacgctcacaatcaattagaagcggtgtgcaaaa 1980
Db 1921 TTTTTCAGCGGACGACCAACACCGCACCTTACATATTTAAACGACCTTGTGTCGCA 1980
Oy 1981 atggaagatctcccaagaagaaatcgtctgagacaaacgatacgtacgcgcacatc 2040
Db 1981 AAAGAGGATTCCTCCGCGGGAAATCGTGGGACAAACGATGATCAACCGCACATTT 2040
Oy 2041 aagcggaacacttccatctcaagggcgagaacggtgttctccgcaatgtttccaaa 2100
Db 2041 AAAGCGGAAATCTTCAAAATTAAGCGGACAGGGGGGTTCCCAATGTTGCAAA 2100
Oy 2101 gtggaagcgatctgacatctaaacacatcagcccaagcagttctcgtctgcacgcat 2160
Db 2101 GTGAAAGCGGATTTGCGCATTTGACCAATCCGCAACGAGTTTGGTTCGCAACGCTAT 2160
Oy 2161 caaagcacacacatctgacacgctcgcagctggaacggtctgacaagctgtacgaanaa 2220
Db 2161 CAAAGCCACACATCTGTACACGTTGCGACCTGAGACGGGTCTGACAAATGTGTCGCAAAA 2220
Oy 2221 accatccagacgataaagtgatgtctcattgagacgaacgacacacagaaggaatgct 2280
Db 2221 ACCATTCAGGACGATTAAGTATTCCTTCAATTTGACCTAGACGACATCAGGCAATGTC 2280
Oy 2281 agccttgccgacagcgtcctaatctaaatctcaaggaacttgcaacactcaacggaatct 2340
Db 2281 GATCTTCCGATCCAGCTCCTCATTTAATCTCAACAGGCTTGCACACATCAACGCGCAATCTT 2340
Oy 2341 atgtagcgagacagacacatatacgtctacgcgcgaacgcgcacacacacacacacac 2400
Db 2341 AGTCGCAATGGCGATACACAGTTATACGTATACGTACAGCACACGCGCACCAAAAGCGCAACTT 2400
Oy 2401 agcctcgtgagcaatgagcccaagcaacatcttaacgaacacataaagcggaacacgctg 2460
Db 2401 AGCCTGAGGCAATGCGCAACGCAACATTTATTAAGCAACATTTAAAGGCAACATGTC 2460
Oy 2461 gctcggacacatgctcattcaatctaaatcaacacacacgctcgaacaaagcgagctgagc 2520
Db 2461 GCTTCGGGCAATGCTTCAATTAATCTAAGCGCACGCGCTGCAACAAAGCGGAGCTTACG 2520
Oy 2521 ctttcgcgaacacgtaagcaaacgtaagccatctcgcacacacacacacacacacacac 2580
Db 2521 CTTTCGGGCAACGCTAAGCAAAAGCTAAGCCATTCGCACTCAACGGAATGCTTCCCTTA 2580
Oy 2581 ggcgataagcgatctcattcaatctaaacacacacacacacacacacacacacacacac 2640
Db 2581 GCGGATTAAGGCAATGCTTCAATTAATTAAGCAACGCTTACCGGACCAATTAAGCGGCG 2640
Oy 2641 aaggaatcgacatctacacttaaaagacagcgaatgagcgtgctgagcagcgaatla 2700
Db 2641 MAGATTAAGGCAATTAATTAATTAAGCAACGCAATGAGCGCTGCTCAAGCGCAATTTA 2700
Oy 2701 ggcacatctaacacacacacacacacacacacacacacacacacacacacacacacacac 2760
Db 2701 GGCATTTTAACCTTGAACACGCAACCATTAACATCAATTCGCGCTATCCGACGATGCG 2760
Oy 2761 gcaagcgcgaaacgagcgagctgagcagatgagcgagcgcgcgcgcgcgcgcgcgcgcgc 2811
Db 2761 GCAAGGCGCAACGCGGAGTGCAGATGAGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 2811
Oy 2812 cgttcctctatctcgtctacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2871
Db 2812 CGTTCCCTATTAATCCCTTACACCGCAACCTTGTGATGAATCCCGCTTCAACACGCTACG 2880
Oy 2872 gtaaacggaacatctgaacgctcaggaagaacatctcgtctctatctgctgaacacacac 2931
Db 2881 GTAAAGCGCAATTTGAAGCGTCAAGGGAACATTCGCTTTATGTCGGAATCTTGGCTAC 2940
Oy 2932 cgcagcgagcaatctgaacgctcaggaagaacatctcgaacacacacacacacacacacac 2991
Db 2941 CGCAGCGCAATTTGAAGCTGGCGGAAAGTTCCGAAAGCACTTACCTTGGCGGTACAC 3000
```

[illegible]

QY	4072	gaacacgaatagcaaaacgtcaaatatcgcacacccggcgcttgatccaacgcgtccgcg	4131
Db	4039	GATTACCGGTACGAAACGTCATATTCGCCACCCCGCGCTTGCAATTCACACCGCTACCGG	4098
QY	4132	gcgggcgaatgaacgaattatcattcaacacggcggaacacatttcacgaagccttat	4191
Db	4099	GCGGGCAATTAGGCAAGTATTTTCATTCAACCGGGCAACACATTTCATCAGCCCTTAA	4158
QY	4192	ttgaagcgtctctataccgagatgcgccttcggcgcaagtcgcgaacgcgcgtcataccgc	4251
Db	4159	TTGAGCTCTGTCATATACCATATGCGCTTTGGGCAAAAGTCCGACACCGCTGTAATACGGCC	4218
QY	4252	gtatctgcgaaggaatttcggcaaaacccgcagtcgcggaatgaggcggtlaaacgcgcgaatc	4311
Db	4219	GTATTGGCTCAGATTTTCGGCAAAACCCGCAATGCGGATAGGCGGCTTAAACGCGGAATTC	4278
QY	4312	aaaggttcacgctgcgcctcccaacgctgcgcgcgcgaagggcgcgcaattggaaagcgag	4371
Db	4279	AAAGGTTTACAGCTGTGTCCTCCACGCGTGGCCGCGCAAAAGGCCCGCAACTGTGAAGGGCAA	4338
QY	4372	cacagcgcgagatcaaatatgaagctacgcgtgtgtaa	4407
Db	4339	CACAGCGCGGCAATCAATTAAGCTACCGCTGTAA	4374
RESULT	3		
AX044007	AX044007	4374 bp	DNA
LOCUS	Sequence 86 from Patent WO0066791.		linear
DEFINITION	AX044007		PAT 24-NOV-2006
ACCESSION	AX044007		
VERSION	AX044007.1	GI:11342893	
KEYWORDS			
SOURCE	Neisseria meningitidis.		
ORGANISM	Neisseria meningitidis.		
REFERENCE	1 (bases 1 to 4374)		
AUTHORS	Pizzetti, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J.C.,		
TITLE	Maignan, V., Galeotti, C., Mora, M., Rattl, G., Scarselli, M.,		
JOURNAL	Scariato, V., Rappelli, R., Frazer, C.M. and Grandi, G.		
FEATURES	Neisseria genomic sequences and methods of their use		
SOURCE	Patent: WO 0066791-A 86 09-NOV-2000;		
BASE COUNT	CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)		
ORIGIN	1..4374		
	/organism="Neisseria meningitidis"		
	/db_xref="taxon:487"		
	1266 a 1163 c 1066 g 879 t		
Query Match	89.7%	Score 3954.8;	DB 6; Length 4374;
Best Local Similarity	94.0%;	Pred. No. 0;	
Matches 4153; Conservative	0;	Mismatches 212;	Indels 51; Gaps 2.
QY	1	atgaatacaacgacgaacgacgaacgaacacacacgcgcaaaagcccttaaacgcgcgcg	60
Db	1	ATGAAAAACAACCGACAAACGACAAACCGAAACACACCGCAAAAGCCCGAAACCGGCGCG	60
QY	61	atccgccttcgcgcgcgccttacttaagccataatgctgtgtgttcgagcattctgcccgaagcc	120
Db	61	ATCCGCTTCTCGCGCTCTACTACTAGCATATGCGCTGTCGGCATTTCTCCCAAGCC	120
QY	121	cgagcggaacacacctatttcggactcaaacacacatacctaactgcgcgcttcgcgaagaat	180
Db	121	TGGCGGGACACACTTATTTCGGCATCAACTCCAACTCATCTGCGCATTTGGCGAAAT	180
QY	181	aaagcaagatctgcatatcgggcgcaaaagataatgaattacaaacaaaaggagagattg	240
Db	181	AAAGCAAGATTTCAATCGCGCGGAAAGATATTAAGGTTTCAACACAAAAGGGGAGATTG	240
QY	241	gtcggcaaatcgatcgcaaaagcccgatgatgaattttctgtgtgatactgcgtaacgacg	300
Db	241	gtcggcaaatcgatcgcaaaagcccgatgatgaattttctgtgtgatactgcgtaacgacg	300

```
Db 241 GTGCGCAATTCATGACAAAGCCCCGATGATTTTCTGTGGTGGCTGACGGC 300
QY 301 gtggcgcaatctggcgcgatcaataatctgtagcgctgacataaaccgagctaac 360
Db 301 gtggcgcgcaatctggcgcgatcaataatctgtagcgctgacataaaccgagctaac 360
QY 361 aatgtctatcttgcgtaggaagcaaccgacatccgacatccgacatccgacatcc 420
Db 361 AACGTTGATTTTGTGGGGAAGGAATATCCGATCAATCTTTTACTTAAATTT 420
QY 421 gtgaagaagaataatataaagaaggaatccgacatccgacatccgacatccgac 480
Db 421 GTGGAAGCAATTAATTAATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 481 atgcgcgcttgcacaaatctgacagatgcagaaaccgttgagatgcagcaatata 540
Db 481 ATGCGCGCTTGTGATTAATTTGTGACAGATGACAGAACTGTTGAATGACGATTA 540
QY 541 gctgagtgagaaatacgcctgattataataaccctgactgctgtctgacatcgag 600
Db 541 GATGGGCGGAATATATGATCAAAATTAATTAATTAATTAATTAATTAATTAAT 600
QY 601 agacataatctggcgctgacatgaagaacgaacgaacgaacgaacgaacgaacga 660
Db 601 AGGCAATATTTGGCGATCTGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 gcaagcgacatcttgcgctgcgctgacatccatccatccatccatccatccatcc 720
Db 661 GCAGAGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 ggcacagtcacatctgtagtaggaagaagaagaagaagaagaagaagaagaaga 780
Db 721 GGCACAGTCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
QY 781 ggaagcgatcttggagaaagtgagtcacacacacacacacacacacacacacac 840
Db 781 GGAAGCGCTCTTTGGCGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 841 tggtaataatctgtagtagtagtagtagtagtagtagtagtagtagtagtagtag 900
Db 841 TGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
QY 901 cagctagtcgtaggaagtgtagtagtagtagtagtagtagtagtagtagtagtag 960
Db 901 CAGCTAGTCGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
QY 961 tcttaagaacacacacacacacacacacacacacacacacacacacacacacac 1020
Db 961 TCTTAAGACACACACACACACACACACACACACACACACACACACACACACAC 1020
QY 1021 aaatcgaatgcacacacacacacacacacacacacacacacacacacacacac 1080
Db 1021 AAATCGAATGCACACACACACACACACACACACACACACACACACACACACAC 1080
QY 1081 caatgtcttaatcttcttcttcttcttcttcttcttcttcttcttcttcttct 1140
Db 1081 CAATGTCTTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1140
QY 1141 ggggtcaacgactacgacacacacacacacacacacacacacacacacacacac 1200
Db 1141 GGGGTCAACGACTACGACACACACACACACACACACACACACACACACACACAC 1200
QY 1201 ggaagaagtgaatctgatacttaccagacacacacacacacacacacacacac 1260
Db 1201 GGAAGAAGTGAATCTGATACCTACGACACACACACACACACACACACACACAC 1260
QY 1261 gaaggaacttctacgctcgccttaaaacacacacacacacacacacacacacac 1320
Db 1261 GAAGGAACCTTCTACGCTCGCCTTAAACACACACACACACACACACACACACAC 1320
QY 1321 atcagatgaggaagtagccttacttggaagtaaacgagcgagcgagcgagcgag 1380
Db 1321 ATCAGTAGGAGTAGCCTTACTTGTGAAAGTAAGGAGGAGGAGGAGGAGGAGGAG 1380
```

```
QY 1381 aaatcggcaaggaacgctgctgcttcaagccaaggggaaacacacacacacac 1440
Db 1381 AAATCGGCAAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
QY 1441 gtggcgcgcaatctggcgcgatcaataatctgtagcgctgacataaaccgagct 1500
Db 1441 GTGGCGCGCAATCTGGCGCGATCAATAATCTGTGAGCGGCGACATTAACGGCG 1500
QY 1501 tttagtaaatctgcttgcgtagcagcagcagcagcagcagcagcagcagcagc 1560
Db 1501 TTTAGTAATCTGCTTGCATGCGCAGCGCAGCGCAGCGCAGCGCAGCGCAGCGC 1560
QY 1561 ttaacacacacacacacacacacacacacacacacacacacacacacacacac 1620
Db 1561 TTAACACACACACACACACACACACACACACACACACACACACACACACACAC 1620
QY 1621 tgcgcttgcctccacgcgacatccacacacacacacacacacacacacacacac 1680
Db 1621 TCGCTTGCCTCCACGCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 1680
QY 1681 caagcaagaatccacgcgacatccacacacacacacacacacacacacacacac 1740
Db 1681 CAAAGCAAGAATCCACGCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 1740
QY 1741 aacaaacacacacacacacacacacacacacacacacacacacacacacacac 1800
Db 1741 AACAAACACACACACACACACACACACACACACACACACACACACACACACAC 1800
QY 1801 gcaacaaacacacacacacacacacacacacacacacacacacacacacacac 1860
Db 1801 GCAACAAACACACACACACACACACACACACACACACACACACACACACACAC 1860
QY 1861 ttaacgcttgcgtaggaagtgtagtagtagtagtagtagtagtagtagtagtag 1920
Db 1861 TTAACGCTTGCCTGCGCAGCAATTAATTAATTAATTAATTAATTAATTAATTA 1920
QY 1921 ttttcaagcgagacacacacacacacacacacacacacacacacacacacacac 1980
Db 1921 TTTTCAAGCGAGACACACACACACACACACACACACACACACACACACACACAC 1980
QY 1981 atcgaagtagtccacacacacacacacacacacacacacacacacacacacac 2040
Db 1981 ATCGAAGTAGTCCACACACACACACACACACACACACACACACACACACACAC 2040
QY 2041 aaagcgaaacacacacacacacacacacacacacacacacacacacacacac 2100
Db 2041 AAAGCGAAACACACACACACACACACACACACACACACACACACACACACAC 2100
QY 2101 gtggaagcgatctgacatcttaagcaatccacacacacacacacacacacac 2160
Db 2101 GTGGAAGCGATCTGACATCTTAAGCAATCCACACACACACACACACACACAC 2160
QY 2161 caaagcacaacacacacacacacacacacacacacacacacacacacacacac 2220
Db 2161 CAAAGCACAACACACACACACACACACACACACACACACACACACACACACAC 2220
QY 2221 accattaccagcagatgaagtgatcttcaatctgagcaagcagacacacacac 2280
Db 2221 ACCATTACCAGCAGATGAAGTGATCTTCAATCTGAGCAAGCAGACACACACAC 2280
QY 2281 agccttgacgacacacacacacacacacacacacacacacacacacacacac 2340
Db 2281 AGCCTTGACGACACACACACACACACACACACACACACACACACACACACAC 2340
QY 2341 agtgcagcgaggaacacacacacacacacacacacacacacacacacacacac 2400
Db 2341 AGTGCAGCGAGGAACACACACACACACACACACACACACACACACACACACAC 2400
QY 2401 agcctctgtggagcagacacacacacacacacacacacacacacacacacacac 2460
Db 2401 AGCCTCTGTGGAGCAGACACACACACACACACACACACACACACACACACACAC 2460
```


181 aaagcaagtttcagtcgagcgcaaaagatattgagtttaccaaaaaagggatgtg 240
Db AAAGCAAGTTTCAATTCGCGGCGAAAGATATTGAGTTTACCAAAAAAGGGGATGTG 6998
QY 241 gtccgcaaatcgatgaagaaagcccgatgatttcttctgtgtatcgcgtacagc 300
Db GTCCGCAAAATTCATGACAAAGCCCGATGATTGTTTCTGTGTGTCGGTAAACGCG 6938
QY 301 gtgcgagcatctgcgagcgatcaatalatgtgaagcgtgtgcacataagcgcgataac 360
Db GTGCGCGCATTTGGTGGCGATCATATATTGTGAGCGTGCGACATAAACGGCGGCTATAAC 6878
QY 361 aatgtgattttgtgtgaggaagcaatcccgatcacacgcgttcttaccgaatt 420
Db AACGTGATTTTGTGTGAGGAAGAAATCCCGATCAACATGTTTACTTATTAAT 6818
QY 421 gtgaaaagaataataataaagcaggaactaaagccatcttaltgycgagatatacat 480
Db GTGAAAAGCAATTAATTAAGCAGGACTAAAGGCGCATCTTATGCGCGCATTAATCAT 6758
QY 481 atgcgcgtttgcacaaattgtcacagatgcgaacccgtgtgagatgacgagatatac 540
Db ATGCCCGTTTGCATTAATTTGTGCAGATGCAGAACCGTGTGAATGACAGATTATAG 6698
QY 541 gatgggtgaataacgcgtgatttaataataatccctgacgtgttcogaatcgagcagc 600
Db GATGGGCGAATAATATATCATTAATAATTAATCCCTGACCGTGTGATTTGGCGACGCG 6638
QY 601 agacaatatctgcgctgtatgaagacgaacccaataacacgcgaagatcatatcat 660
Db AGCAATATTTGGCGATCATGATGAAGTGAAGCCCAATTAACCGGAAAGTTCAATATCATAT 6578
QY 661 gcaagcagcatattctgtgctgcgtgcgaataacattgagcaaaatgagatgagatgt 720
Db GCAAGCGCATTTCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6518
QY 721 ggcacagcaacttaagtagcgaanaaataacabagccatgatttlttaccaca 780
Db GGACAGTCAACTTAAGTAGTGAATAAATAAATGAAGCCATATGCTTTTACCAACA 6458
QY 781 ggaagctcatcttgagcagcagtgctcaccatgltatctatgagcccaaaagcaaaag 840
Db GAGGCTCATTTGGCGACAGTGGCTCAACCATGTTTATCTATGATGACCCAAAAGCAAAAG 6398
QY 841 tggtaataaaggggattgcaaacgaagcaacccctatatggaanaagcaatgcttc 900
Db TGGTAAATTAATGGSATATTGCAAAAGGCAACCCCTATATAGAAAAGCAATGGCTTC 6338
QY 901 cagctagctcgtaagaattggtctatgatagaatccttgcgtgagatacccatcagta 960
Db CAGCTGCTTCGTAAGATTTGTTCTATGTAATCTTGTGGAGATACCCATTCAGTA 6278
QY 961 tctacgaacccatcaaaaatgggaatacttctttaaagcaataaataatgagcaga 1020
Db TCTACGAACCAAGTGAATGGGAATACCTCTTTAAAGCAGTATATATATGACACAGTA 6218
QY 1021 aaatcgaatgcgaataataacacattctctacattatagattaaacacgaacgctt 1080
Db AAATCGAATGCGAATAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 6158
QY 1081 caattgttaatgltcttcttctacagacagacgaagaaacgctgtatcatgctcaggt 1140
Db CAATTGTTAAATGTTCTTATTCGAGACAGACAGAACGCTGTTATCATCTGCAAGT 6098
QY 1141 ggggtcaacaattatcgacagcagcagatgaataatgagaanaaataattctcttcttgc 1200
Db GGGGTCAACAATTATCGACCCAGCATGAATTAATGGAATAATTAATTAATTAATTAATTAATTA 6038
QY 1201 ggaagaaagtgaattgatacttaaccagcaacatcaaacgaagcgcggtgtgtatct 1260
Db GGAAGAAAGTGAATTGATCTTACAGCAACATCAATCAAGGTGCTGAGGATTAATATTC 5978

QY 1261 gaggtaattttagcgtctgcctaaacacgaacgctgagcagcggttcatt 1320
Db CAAGAGATTTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5977
QY 1321 atcagtgatgcatccgttacttgaagaataacgagcgtgcaaacgacgctgtcc 1380
Db ATCAGTGAAGCAATACCTTACTTGAATAAACCGCGCTGCGCAACCGCTGCTTC 5858
QY 1381 aaatcgcgaagcagcgtctgttcaaacgaaggggaaacccaagcgtctgcagc 1440
Db AAATCGGCAAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5857
QY 1441 gtggcgagcgttaagtcaltatgatacagcagcgagatcaagcgaacaaacagc 1500
Db GTGGCGAGGCTACGATATTGTGATGACAGCAACATTAAGGCAAAAACAAGCC 5738
QY 1501 tttagtaatcgcgttgtltagcagcagggagcgtgcaactgaatgccaatacag 1560
Db TTTAGTGAATCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5737
QY 1561 ttcaaccccgcaaacctctatcttgcgttgcgagcaggtttgagttgaaacgagcat 1620
Db TTCAACCCCGCAAACTCTATTTGCGCTTTCGCGCGAGCTTTGATTTAAACGGCAT 5618
QY 1621 tgcgttcgttccacgcgcatcaaaatacagatgaagggcgatgattgtcaaccat 1680
Db TCGCTTCGTTCCACCGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5558
QY 1681 caagacaagaatccacgcttaccatlaacgcaataaagatatlaacaaacggaat 1740
Db CAAGCAAGAAATCCACGTTTACCATTAACAGGCAATTAAGTATTTGCTACACCGCAAT 5498
QY 1741 acaacaacgttgatgaacaaagaagaatgtctacacagcgttgtgttgcgagaaagat 1800
Db AACACAGCTTGATGACAAAAGAAATTCCTTACACGCTTGCTTGGCGAAGAAAGAT 5438
QY 1801 gcaacaaacgaacgagcgtctcaatcgaatlaaccaacggaagaagcagatcgact 1860
Db ACAGCAAAAGAAAGGCGCGCTCACCTTGTATTAACGCGCGCGCGAGAAACCGCGAC 5378
QY 1861 ttactgcttccgagcgaacaaatttaacgcaatctccgcaacaaacgcaacgct 1920
Db CTGCTGCTTCGCGCGGAAATTTAAACGGCAACATTAACGCAAAACGCGCAAACTG 5318
QY 1921 ttcttcagcagcgaacggaacgcaacgcaacatatttgaagaagcggtgtgcaaa 1980
Db TTTTTCAGCGGCAACCGCACCGACGCTACATATTTAAACGACACTTGGTGCACA 5258
QY 1981 atggaagatcccaacgaaggaatcgtgtggaacaaagattgagtcgacgcaatt 2040
Db AAAGAGGCGATTCCTCGCGGGGAATGCTGTGGACAAACGACTGGATCAACCGCATTT 5198
QY 2041 aaagcggaacactccataltcaagcggaacagcgtgttcccgcaatgttgcga 2100
Db AAAGCGGAACCTTCCAATTTAAAGGCGGACAGCGGTGCTTCCGCATATGTTGCCAA 5138
QY 2101 gtggaagcgaatgagcatlaagcaatcaacgcccagcagtttgcgtgcacgcat 2160
Db GTGAAGGCGATTTGCAATTTGCAATTCACGCCCAACAGTTTGTGTGCGACCGCAT 5078
QY 2161 caaagcacaacatcttlaacgcttcgagcgtgagcgtgtgcaaaattgacggaana 2220
Db CAAAGCCGCAACATCTGACGCTGAGCTGAGCGGCTGTGCAAAATTTGTCGAAAA 5018
QY 2221 accattacgcagataagatgattgtcttcttgaagcagcagatcagaagcaatgtc 2280
Db ACCATTACCGCATTAAGATGATTTGCTTATTAAGACCGACATCAACCGCAATATGTC 4958
QY 2281 agccttgcagcagcgtcttlaaatcacaaggaattccaacactcaacgcaactc 2340
Db GATCTTGCGCATCAGCTCATTTAAATTCACAGAGGCTGCCACACTCAACGCGCAATCTT 4898
QY 2341 agtgcagcggaagacagcaactatagtgtaagcgaacgcaacccaacgcaactc 2400

4897 AGTGCATAATGGGATACACGTTATATACATTCAGCCCAACAGCGCACCAAAAGCGAACCTT 4838
4401 agctctcgtaggaatgcccagaacaatttaataagccacattaaacgagcacatcg 2460
4837 ACCCTGCTGGGCAATGCGCAAGCAACATTATCAAGCCCAATTAAAGGCAACATATCG 4778
2461 gcttcgagcaatgctcatttaataagcaacaacgcccgtacaaaacgagctacg 2520
4777 GCTTGGGCAATGCTTCAATTAATTAAGCGCACCGCCGTACAAAACGGCAGCTGACG 4718
2521 ctttcgagcaatgcttaagcaacaacgtaagccattccgactccgaacggaatgtcccta 2560
4717 ctttcgagcaatgcttaagcaacaacgtaagccattccgactccgaacggaatgtcccta 4658
2581 gccgataagagcagatltccatttgaanaacgagccgttaccggaataaactcagcgagc 2640
4657 GCGGATTAAGGCAATGCTTCAATTAATTAAGCGCACCGCCGTACAAAACGGCAGCTGACG 4598
2641 aaggaatacgcatacacttaaaagacagcagatgagcgttcgagcagcagatla 2700
4597 AAGGATTAAGGCAATGCTTCAATTAATTAAGCGCACCGCCGTACAAAACGGCAGCTGACG 4538
2701 ggcgaatlaaaccttgagcaaacgcccatacactcaattccgactatcgacagatcg 2760
4537 GGCATTTTAACCTTGACACCCACCATTAACCTTAACCTTAACCTTAACCTTAACCTTAAC 4478
2761 gcaagcgagcaaacgagcagtgagcagatgagcagcagcagcagcagcagcagcagcagc 2811
4477 GCGAGGCGGCAAAACCGCGAGTGCAGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4418
2812 cgttcctcattatcgttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2871
4417 CGTTCCTATTTCGCTTACACCGCCCACTTGTGAGATCCCGTTTCAACACCTGACG 4358
2872 gtaaacgcaaatgcaaggttcaaggaacatccgcttattgtcggaacttccgctac 2931
4357 GTTAAAGGCAAAATTAACCTTGACACCCACCATTAACCTTAACCTTAACCTTAACCTTAAC 4298
2932 cgcagcgagcaaatgcaaggttgcaggaaggttcggaagcagcagcagcagcagcagcagcagc 2991
4297 CGGAGGCAAAATTAACCTTGACACCCACCATTAACCTTAACCTTAACCTTAACCTTAAC 4238
2992 aatacgcggaacgaacccgtaagttcgaagcagcagcagcagcagcagcagcagcagcagcagc 3051
4237 AATACCGGCAACGAACTTCAAGCTTCAACCTTGACACCCACCATTAACCTTAACCTTAACCTTAAC 4178
3052 aacgcgctgcggaatcttaattcaacctgcaaaacgaacagcagcagcagcagcagcagcagc 3111
4177 AATACCGGCAACGAACTTCAAGCTTCAACCTTGACACCCACCATTAACCTTAACCTTAACCTTAAC 4118
3112 tgcgctatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3171
4117 TGCGCTTACCAACTTCCGCAAGCGCGAGTTCGCGCTGCAATATCCGCTCAAAAGAA 4058
3172 caagaacttccgacaacatcgcaagcgaggaagaaacagagcgccgttgaagcagaa 3231
4057 CAAGACCTTCCGCAACCTTCCGCAAGCGCGAGTTCGCGCTGCAATATCCGCTCAAAAGAA 4023
3232 caagcacaacttgcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3291
4024 -----GCCAAAAACAGCGCGCAAAAGCAACAGCGCAAAAGCGCTTGACGCG 3980
3292 ctagctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3351
3979 CTGATTGCGCGCGCGCGCATGCGCTGCAAAAGCAAGAGCGCTTGCGCAACCGCGCGCG 3920
3352 cagcagcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3411
3919 CAGGCGGCGCGCGCGCAATATGCGCATTAAGCGCGCGAGGAAAGCAAAAGCGGCTGACG 3860
3412 gcgataaagacacgcgcttgcggaacagcagcagcagcagcagcagcagcagcagcagcagcagc 3471

3859 GCGGATTAAGGCAACCGCTTGCGCAAAACAGCGCGAGCGGAAACCGCGCGGCTTACACG 3800
3472 gcttcctcccgagcccgccgagcccgagcagcagcagcagcagcagcagcagcagcagcagcagc 3531
3799 GCTTCCCGCGCGCGCGCGCATGCGCTGCAAAAGCAAGAGCGCTTGCGCAACCGCGCGCG 3740
3532 ccccaacgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3591
3739 CCGCAACCGCGCGCGCGCATGCGCTTACGCGGCAATGCGGCTTGAGTGAATTTTTC 3680
3592 gcaagc 3651
3679 GCGAGCTTAAGGCAACCGCTTGCGCAAAACAGCGCGAGCGGAAACCGCGCGGCTTACACG 3620
3652 cgcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3711
3619 GCGCGCAACCGCGCTTGCGCAAAACAGCGCGAGCGGCAATGCGGCTTGAGTGAATTTTTC 3560
3712 ttcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3771
3559 TTCGCGCGCTTACGCGCGCAACCGCTTGCGCAAAACAGCGCGAGCGGCAATGCGGCTTGAGTGAATTTTTC 3500
3772 agcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3831
3499 AGCGGCGCGCTGCGCATGCGCTTACGCGGCAATGCGGCTTGAGTGAATTTTTC 3440
3832 atcggaactcgcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3891
3439 ATCGGCAACCTGCGCGCGCAACCGCTTGCGCAAAACAGCGCGAGCGGCAATGCGGCTTGAGTGAATTTTTC 3380
3892 ttcgagc 3951
3379 TTTACATCGCGCATGCGCGCGCGGCTTACGCGGCAATGCGGCTTGAGTGAATTTTTC 3320
3952 agaggaacacacgcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4011
3319 GAGGCAAAATCGCGCGCGCGCGCATGCGCTTACGCGGCAATGCGGCTTGAGTGAATTTTTC 3260
4012 ttcgagc 4071
3259 TTTGCGCGGATTTGCGCGCGCGCGCATGCGCTTACGCGGCAATGCGGCTTGAGTGAATTTTTC 3200
4072 gattacagc 4131
3199 GATTACCGCTTACGCGCGCGCGCATGCGCTTACGCGGCAATGCGGCTTGAGTGAATTTTTC 3140
4132 ggcgagc 4191
3139 GCGGCAATTAAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3080
4192 ttcgagc 4251
3079 TTTGAGCTTCTGCTTACGCGCGCGCGCATGCGCTTACGCGGCAATGCGGCTTGAGTGAATTTTTC 3020
4252 gttatgagc 4311
3019 GTATTGCTTACGCGCGCGCGCATGCGCTTACGCGGCAATGCGGCTTGAGTGAATTTTTC 2960
4312 aagagcag 4371
2959 AAGGCTTACGCGCGCGCGCATGCGCTTACGCGGCAATGCGGCTTGAGTGAATTTTTC 2900
4372 cagc 4407
2899 CAGAGCGCGCGCGCATGCGCTTACGCGGCAATGCGGCTTGAGTGAATTTTTC 2864

RESULT 5
AX044034/c AX044034 349860 bp DNA linear PAT 24-NOV-2000
LOCUS
DEFINITION Sequence 113 from Patent WO006791.
ACCESSION AX044034
VERSION AX044034.1 GI:11342918

OY	1741	aacacaaacttgatagcaaaaagaagaatggctacaaaggttggtcttgagcaagaagt	1800
Db	287451	AACACAAAGCTTGATAGCAAAAAGAAATGGCTACACAGCTTGTTGGCCAGAAAGAT	287392
OY	1801	gcaaccaaaacgaacgggagctcaatctgaattacaacacggaagaagcgaatgcact	1860
Db	287391	ACGACCAAAACGAACGGCGGCTCAACCTTGTTTACCACCCGCGCCCAAAACGCGACAC	287332
OY	1861	ttactgcttcctcgcgagaaanaatttaaaacgcaatcatcagcaaaacaaacgcaaatg	1920
Db	287331	CTGCTGCTTTCGGCGGAACAAATTTAAAGGACATCACGCAACCAACAAACGCGAAACTG	287272
OY	1921	ttttcagcgagcagccagacacgcgcgcctaataatcatttgggaagcggtgggtcaaa	1980
Db	287271	TTTTTTCAGCGGACACCAACACCGCACGCTACAAATCATTTTAAACGACCATTTGGTGCA	287212
OY	1981	atggaaggtatcccaacaaggagaatcgtgttgagacaagatltgattcgacgcacattt	2040
Db	287211	AAAGAGGGCATTCCTCCCGGGGAAATCGTGTGGGAACAGCACGTGATCAACCGCATTTT	287152
OY	2041	aaagcggaaacactcccatctcaagggcgagcaagcgggtgtttcccgcaatgtttcgaaa	2100
Db	287151	AAAGCGAAAACTTCCAATTTAAAGGGGACAGCGGTGGTTTCCGCGAATGTGGCAAA	287092
OY	2101	gttgaagaagcatttgacatttaagcaatcaagcccaagcagtttctgtgtcgacccgat	2160
Db	287091	GTGAAAGCCGATTTGGCATTTTGAGCAATACGCCCAACGACTTTTGTGTGCGACCGCAT	287032
OY	2161	caaaacacacaaatctgtacacgttcgacatgagcgggtctgacaagtgtlacagcaaaa	2220
Db	287031	CAAAAGCCACAAATCTGTACACGTTCGACGTGGACCGGTCTGTCAAAATGTGTGCAAAAA	286972
OY	2221	accattaccgcgcaataagatgtgtctcatctagcaagaacacatcagaagcaatgtc	2280
Db	286971	ACCATTTACCGACGATTAAGATGTGTTGCTTCAATGACTAAGACGCAACATACGCGCATATGTC	286912
OY	2281	agccttcgcgatacagcttcaatttaaatctcaacgagacttgcacacttaacggaactt	2340
Db	286911	GATCTCGCGATACGCTCATTTTAAATCTCACAGGGTGTCTCCACACTCAACGGCAATCTT	286852
OY	2341	atgtcagcgcggaagacacgacatctacgtctcagcgcaacgcaccccaaaaacgcaactc	2400
Db	286851	AGTGCATATGGCGATACACGTTTATACGTACAGCCACACACCCCAACCAAAACGGCAACCTT	286792
OY	2401	agccctcggtggcaatgtcccaagcaacacattcaatcaaaacacattaaacgcaacacatcg	2460
Db	286791	AGCCCTCGTGGGCATGTGCCCAAGCAACATTTATATCAACCCACATTAAACGCAACACATG	286732
OY	2461	gcttcgcgaacatgctcaattatctaaagacaacagcgctgacaaaaacggtcgtcgaag	2520
Db	286731	GCTTGGGGCAATGTCTCATTTTAAATCTTAAGGACACCCCGCTACAAAACGGCAATGTGACG	286672
OY	2521	ctttccgcacaacgcgtaaggcaaaacgtaagacatctcgcgactcaaaacgcaatgtctcccta	2580
Db	286671	CTTTCGCGCAACGCTTAAGGCAAAAGTAAAGCATTTCCGCACTCAACGGTATGTCTCCCTA	286612
OY	2581	ggcgaataaggaagatcatcttctgaaaaaacgcgcgtcttaccggaanaaatcagcgcgcg	2640
Db	286611	GCCGATTAAGGCAATTTTCATTTTGAAGACAGCCGCTTTTACCGGACAAATTCACGGCGG	286552
OY	2641	aaggaatacgcgcatatacatttaaaagaacaggaatlgacgtctgcgtctggcgacagatta	2700
Db	286551	AAGGTATACGCGATTAACATTTAAAGACAGCGATATGACGCTGCCGTAAGCACGGAATTA	286492
OY	2701	ggcaatttaaaccttgacaacgcgcacacattacacttaattccgctctctgaaacgaatgg	2760
Db	286491	GGCATTTTAAACCTTGAAACAAAGCCACCATTTACATTCATTTCCGCTATTCGCCACGATGCG	286432
OY	2761	gcaagcgcgcaaacccgcgaatgtcgagcagaatgcgcgcg-----gcgcgttcgcgc	2811
Db	286431	GCAGGGCGCGCAAAACCGGAGTGCACACATATCGCCGCGCGCTTCGCGCGCTTTCGCGC	286372

[illegible]

[illegible]

Db	1501	TTTGTGTAATTCGGCTTGGTCACGGGCGAGGGGAGACGGTCAACCTGAATTCGGATTAATCAG	1560
OY	1561	ttcaaccggcgaaacaactctcttccggctttgcgpgcgagacgcttttgatttgaacggcgat	1620
Db	1561	TTTCACCCCCGACAAACTCTATTTCGGGCTTTCGGGGCGAGACGCTTTSGATTTAAACGGGCAT	1620
OY	1621	tgccttccgttccacgcgcattcaaataccgatgtagggcgagatgttgcataaccaat	1680
Db	1621	TGCGTTCCTTGGTTCACCGTATTTCAAAATTAACCGATGAAGAGGGCGCATATTTGTCACCCACAT	1680
OY	1681	caagaacaaagaatccaccggttaccatttaacggcaataaagaatattactacacacggcgcat	1740
Db	1681	CAAGACAAGAAAGATTCACCGTTTACCATTTAACGCAATTAAGATTTGTCTATCAACCGCGCAT	1740
OY	1741	aacacacacttggatagcaaaaagaagaattgcaccacacgcttggcttggcgagaagaat	1800
Db	1741	AACAAACGCTTGGATGATCAAAAAGAAATTCGCTCAACAGTGTGGTTGGGAGAAACAT	1800
OY	1801	gcaaccacaaacgaacggcggttcacatcgaattcccaacgggaagaagcgatctgcacat	1860
Db	1801	ACGACCAAAAGAGACGGGGCGGCTTCACCTTGTTTACCAACCCCGCGAGAAAGCCGACAC	1860
OY	1861	ttactgcttccgcggcggaacaaatttaaacgcaatatacgaacaaacacgcgcaaatg	1920
Db	1861	CTGCTGCTTTCGCCGGGCAACAATTTTAAAGCGCAATACGCCAAACAAACGGCAAACTG	1920
OY	1921	tcttcacggcgagacggacacacgcacgcctaaacatcttgaagaagcgggcttgatcaaa	1980
Db	1921	TTTTTTCAGCGGCACACCCACACACCGCAGCGCTACAAATCATTTTAAAGACCATTGGTGCAG	1980
OY	1981	atgtaagaatccccacaaaggagaatcgtgtgagacaacgatctgagtcgacgcacatt	2040
Db	1981	AAAGAGGCGCATTCCTCGCGGGGAAATCGTGTGGCAACAGCATGGAATCAACGGCGCATTT	2040
OY	2041	aaacgcgaaaaactctcctatccaggcggaacgaagcggtgttcccgcaatggttgcga	2100
Db	2041	AAAGCGGAANAATTCCAATTTAAAGCGGAGACAGCGGCTTTCGCCGATATTTGCCAA	2100
OY	2101	gtgtaagcgcatctggaatttaagcaatcacgcccaagcagtttccgtgtgcacgcgat	2160
Db	2101	GTGAAAGCGCGATTTGGCGATTTTGACCAATTCACGCCCAAGCAAGTTTTGGTGTGCGACCGCAT	2160
OY	2161	caaaagcacaacaatctctatacagcttcgagctggaacgagctctacaagcttgttaccgaa	2220
Db	2161	CAAAAGCCACAAATCTGTACACGTTGCGAGCTGACGCGGTGTGACAAATTTGTGTCCAA	2220
OY	2221	accattaccacgcataaagatgattgtcttcaatctgaagaagaacgcatcagaagcagctc	2280
Db	2221	ACCAATTACCGACGATTAAGATGATTTGCTTCTATTGTACTAABCCGACATCAGCGGCAATGTC	2280
OY	2281	agccttgcgcattcaagcttcaatataatctcaagagccttgcacacactcaagcgcaattc	2340
Db	2281	GATCTTGCCGATTCACGCTCATTTTAAATCTCNCAGGCGTTGCCACACTCAACGGCAGTCTT	2340
OY	2341	agtgacagcggaagacacgacctaiaaggtatacgacgcaacgcccaacaaacggcaactc	2400
Db	2341	AGTCAAAATGGCGATATCACGCTTATACAGTCAAGCCACTACGCCACCCAAACCGCAACTT	2400
OY	2401	agccctgtgggcaatctgccaagaacaacttaatcaagcacactiaaagcgcaacacatcg	2460
Db	2401	AGCCCTGGGGCATGTGCCCAAGCAACATTTAATCAAGCCACATTTAAAGCGCAACACATCG	2460
OY	2461	gcttcgaaacaatgcttaattaactcraagaagaagaacgacgctgaacaaacggcgacttgcg	2520
Db	2461	GCTTTCGGCAATGCTTATTATTCTTAAGGACACACGCCGTACAAAAGCGCAGTGTGACG	2520
OY	2521	ctttccgacaaacgcttaagcaaacgtaagcgaattcgcgaactaaagcgcaatgcttcccta	2580
Db	2521	CTTTTCGGCAACGCTTAAGGCAAAACGTAAGCGATTTTCGCCACTCAACGGTAATGTCTCCCTA	2580
OY	2581	gcgcataagcgagatctcatcttgaanaacgcgcctttaccggaanaatacagcgcgcc	2640
Db	2581	GCCGATTAAGGCGATGATTCATTTTGTGAAGGACGCGCTTACGGGCAATACGCGCGCGC	2640

QY 2641 aaggaatacgtacattacattaaagacgagatgacgctgctgagcaggaatt 2700
DB 2641 AAGGAATACGTAATTACATTAAAGACGAGATGACGCTGCTGAGCAGGAATTA 2700
QY 2701 ggaattaaaccttgacacacccacattacattacattacattacattacatt 2760
DB 2701 GGAATTAAACCTTACACACCCACATTACATTACATTACATTACATTACATT 2760
QY 2761 ggaagcgcaaaacggcgagatgacgagatgacgagatgacgagatgacgag 2811
DB 2761 GGAAGCGCAAAACGGCGAGATGACGAGATGACGAGATGACGAGATGACGAG 2811
QY 2812 gcttcctattatcgcttcgacgacattcgacgacattcgacgacattcgacg 2871
DB 2812 GCTTCCTATTATCGCTTCGACGACATTGACGACATTGACGACATTGACGAC 2871
QY 2872 gtaaacggaatttgacgagatgacgagatgacgagatgacgagatgacgag 2931
DB 2872 GTAACGGGAATTGACGAGATGACGAGATGACGAGATGACGAGATGACGAG 2931
QY 2932 cgaagcgcaaaattgacgagatgacgagatgacgagatgacgagatgacgag 2991
DB 2941 CGAGCGCAAAATTGACGAGATGACGAGATGACGAGATGACGAGATGACGAG 3000
QY 2992 aatacggcaaaacggcgagatgacgagatgacgagatgacgagatgacgag 3051
DB 3001 AATAACGGCAAAACGGCGAGATGACGAGATGACGAGATGACGAGATGACGAG 3060
QY 3052 acacgctgttcggaatttgacgagatgacgagatgacgagatgacgagatgac 3111
DB 3061 AACACGCTGTTCGGAATTGACGAGATGACGAGATGACGAGATGACGAGATG 3120
QY 3112 tggcgttacaagatgacgagatgacgagatgacgagatgacgagatgacgag 3171
DB 3121 TGGCGTTACCAAGATGACGAGATGACGAGATGACGAGATGACGAGATGACG 3180
QY 3172 caaggaatttcggaatttgacgagatgacgagatgacgagatgacgagatg 3231
DB 3181 CAAGGAATTTCGGAATTGACGAGATGACGAGATGACGAGATGACGAGATG 3243
QY 3232 caaggaatttcggaatttgacgagatgacgagatgacgagatgacgagatg 3291
DB 3244 -----GCCAAAAACAGCGGGAAGCAACAGCGGCAAGCGCTTGACGCG 3298
QY 3292 ctgattcgcgcgacgagatgacgagatgacgagatgacgagatgacgagatg 3351
DB 3259 CTGATTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3368
QY 3352 caggaagcgcggaatttgacgagatgacgagatgacgagatgacgagatgac 3411
DB 3319 CAGGACGCGGGAAGATGCTGCGGCAATGCGGCGGGAAGCAACAGCGGCTG 3378
QY 3412 ggggaataaagacgagatgacgagatgacgagatgacgagatgacgagatg 3471
DB 3379 GCGGAATAAGACGAGATGCTGCGGCAATGCGGCGGGAAGCAACAGCGGCT 3438
QY 3472 gcttccttcg 3531
DB 3439 GCTTCCTTCG 3498
QY 3532 cccgaacggcgagatgacgagatgacgagatgacgagatgacgagatgacgag 3591
DB 3499 CCGCAACGGCGAGATGCTGCGGCAATGCGGCGGGAAGCAACAGCGGCTG 3558
QY 3592 ggcgaacgaacgagatgacgagatgacgagatgacgagatgacgagatgacg 3651
DB 3559 GGCACGACGAGATGCTGCGGCAATGCGGCGGGAAGCAACAGCGGCTG 3618
QY 3652 cgcgaacggcgagatgacgagatgacgagatgacgagatgacgagatgacgag 3711
DB 3619 CGCGACGAGATGCTGCGGCAATGCGGCGGGAAGCAACAGCGGCTG 3678

QY 3712 ttcgcgcttaccgacgaacacacgacgtgacgaacatcgatgacgaacacacgtgc 3771
DB 3679 TTCCGCGCTACCGCAACAAACGACGCTGCGCAATCGATGACGAAACCTCGCG 3738
QY 3772 agcgcgcgctgacgacacgtgacgaacacacgacgtgacgaacacacgtgc 3831
DB 3739 AGCGGCGCTACCGCAACAAACGACGCTGCGCAATCGATGACGAAACCTCGCG 3798
QY 3832 atcggaacacacgacgacgtgacgaacacacgacgtgacgaacacacgtgc 3891
DB 3799 ATCGGAACCTGCGGACGCGCTGCGCAACGCGGCAATCGATGACGAAAC 3858
QY 3892 ttcgaacacacgacgacgtgacgaacacacgacgtgacgaacacacgtgc 3951
DB 3859 TTCTACATCGGATGACGCGGCGGCGGCTGCGCAATCGATGACGAAACCT 3918
QY 3952 agagcaaaacacgacgacgtgacgaacacacgacgtgacgaacacacgtgc 4011
DB 3919 GAGGCAAAATCGCGCGCGGCGGCTGCGCAATCGATGACGAAACCTCGCG 3978
QY 4012 ttcggaacacacgacgacgtgacgaacacacgacgtgacgaacacacgtgc 4071
DB 3979 TTGCGGCGATTCGCGGACGCGCTGCGCAACGCGGCAATCGATGACGAAAC 4038
QY 4072 gattacgaataaagacgacgtgacgaacacacgacgtgacgaacacacgtgc 4131
DB 4039 GATTACCGCTACGAAACGCTCAATGCGCACCGCGGCTGCGCAATCGATG 4098
QY 4132 ggcggaataaagacgacgtgacgaacacacgacgtgacgaacacacgtgc 4191
DB 4099 GCGGCGATTAAGCGGCAATGATGATGATGATGATGATGATGATGATGATG 4158
QY 4192 ttcgaacacacgacgacgtgacgaacacacgacgtgacgaacacacgtgc 4251
DB 4159 TTGACCGCTGCTTACGCGGCGGCGGCTGCGCAACGCGGCAATCGATGAC 4218
QY 4252 gattacgaataaagacgacgtgacgaacacacgacgtgacgaacacacgtgc 4311
DB 4219 GATTACCGCTGCGGATGCGGCAATGCGCACCGCGGCTGCGCAATCGAT 4278
QY 4312 aaaggttcaacgctgacgacgtgacgaacacacgacgtgacgaacacacgtgc 4371
DB 4279 AAAGGTTTACGCGGCTGCGGCAATGCGCACCGCGGCAACGCGGCTG 4338
QY 4372 cacaagcgcgacgaataaagacgtgacgaacacacgacgtgacgaacacacgtgc 4407
DB 4339 CACAGCGGCGGCAATGATGATGATGATGATGATGATGATGATGATG 4374

RESULT 7
NME296277 4386 bp DNA linear BCT 14-JAN-2002
LOCUS Neisseria meningitidis app gene for adhesion penetration protein,
DEFINITION strain 24181.
ACCESSION AJ296277.1 GI:11071864
VERSION AJ296277.1 GI:11071864
KEYWORDS adhesion penetration protein; App gene.
SOURCE Neisseria meningitidis.
ORGANISM Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.

REFERENCE
1 (sites)
AUTHORS Hadji,H.A., Wooldridge,K.G., Robinson,K. and Ala'Aideen,D.A.
IDENTIFICATION and characterization of App: an immunogenic
autotransporter protein of Neisseria meningitidis
Mol. Microbiol. 41 (3), 611-623 (2001)
2 (bases 1 to 4386)
REFERENCE
AUTHORS Ala-Aideen,D.A.
TITLE Direct Submission
JOURNAL Submitted (27-OCT-2000) Ala-Aideen D.A., Microbiology, University
of Nottingham, Molecular Bacteriology and Immunology Group,
University Hospital, Nottingham NG7 2UH, UNITED KINGDOM

[illegible]

QY	1321	atcagtcgtgcacgtacacgttactctggaaagtaaacgctgtgcacaacgacgctgtcc	1387
Db	1294	ATCAGTGAAGACAGTACCGTTACTTGTGAANGTAAACCGCGTGGCAACACGCGCTGTCC	1355
QY	1381	aaatccgcaaaagccagcgtgtcgtgttccaagccaaaggggaaacccaagctccgttcacg	1440
Db	1354	AAAATCGCAAAAGCGCGCTGCACGTTTCAAGCCAAAGGGGAAACCAACGAGCTGCATACG	1413
QY	1441	gtcggcgcgcgtlaaagtcacatttagtaccgacgacgcagatcaagccaaaaaacaaagcc	1500
Db	1414	GTGGCGCGACGGTAAAGTCATTTTGGATCACAGGACAGCGTAAAGGCCAAAAACAAGCC	1473
QY	1501	tttagtgaatccggcttgggtccagcggaagggagacggttccaactgaaatccgataatcag	1566
Db	1474	TTTACTGAATCGCGCTTGTCACGCGGAGGGGTTACGGTGCMACTGAATCCGATATATCAG	1533
QY	1561	ltcaaccccgcaaaactctatatttcgcttccgcgcgacgtttgatttgaacgggcat	1620
Db	1534	TTCAACCCCGCAAAACCTATTTCGGCTTTCGGGGGAGGATTTTGGATTAAACGGGGAT	1593
QY	1621	tcgcttgcgttccacgcgcgatccaataaccgaatgaaagggcgatgtatccaaccaat	1680
Db	1594	TGCGTTTGCTTCCACCGGATTCAAAATATCCAGATGAAAGGGGCGATGTATGTCAATCATAT	1653
QY	1681	caagacaagaatccaccggttaccatctacagcgcaataaagatatattactcaacgcgc---	1737
Db	1654	GCCACACAAACATCCACCGTTCACATTACAGGGAGATGAAGATTTACACACACCGAGTGT	1713
QY	1738	---aatacaacaactctgatatgcaaaaaaagaattctgcacaaagcttgtgtttgcag	1794
Db	1714	AAGATATTCATTAACACTTAATTTACAGAAAGAAATTTGCTTACACGCTGTGTGGCCAG	1773
QY	1795	aaagaatgacaacaaacgaacgcgcgcgtcaatctgaattacaacccggaagaacgcgat	1854
Db	1774	AAAGATAGACCAAAAGCAAGGAGGCGGCTCAACTTGTGTACAGCCGCCGACAGAAC	1833
QY	1855	cgcacttaccgcttcccgcgcggaacaatttaaacggcaatataccgnaaaacaagcc	1914
Db	1834	CGACCCGCTGCTGCTTCCGGGGAAACAAATTTAAACGCAACCTCACGGAAACAAACGCG	1893
QY	1915	aaactgttlttcacgcgcgaacgcgaacgcgcacgcctacaatcaatctaaagacggcggtg	1974
Db	1894	AAACGTGTTTTACCGGCAACACGACACCGCACCGCTCAATCATTTAGAGCGGGGCG	1953
QY	1975	tcaaaaatgggaaggtatcccaacaaggaagaaatcgtgtggaaacaagattggatgcacgc	2034
Db	1954	TCAAAAATGGAAGGTATCCCAACAAGAGAAATGCTGTGGGACAAACGATCGATCAACGCC	2013
QY	2035	acatttaagcggaaaaacttccatlatctcaggcgcaagcgggtgttcccgcaatgtc	2094
Db	2014	ACGTTTAAAGGGAATAATTTTCCATTATTTACAGGGGAGGCAAGCGGTCAATTTCCGCAATGTT	2073
QY	2095	gcacaagatgggaagcgcatctggcatcttaagcaacaacgcgcgaacgcagtttcggtgcga	2154
Db	2074	GCCAAAGGGAAGGCGATTTGGCATTTGAGCAATCAACGCCCAAGCATGTTTTGTGTCCGA	2133
QY	2155	ccgatacaaacgcacaaactcttaaacgcttggagatctgaacggcggtctcaacaagttgtac	2214
Db	2134	CCGCATCAAAACCATTAACATCTGTACACGTTCCGATCTGACCGGCTGTGACAAATTTGTCT	2193
QY	2215	gaaaaaaccatctacgcgcagataaagtgtatgtctcatcttagcaagaacgcgacatcaagag	2274
Db	2194	GAATAAACCATTTACCGACGATGAAGATGATGCTTCATTAACATCAACGACATCCAGCGCG	2253
QY	2275	aatgtcagccttgcgcatacgcctcatcttaaatctcaacggagatttgcacacactcaagcc	2334
Db	2254	AATGTCACCGCTTGGCGATTCACGCTCATTTAAATCTCACAGGCGCTTGCCACACTCAACGCG	2313
QY	2335	aattcttgtagcgcggaagacgcgcacitaaagttaggtttagcgcgaacgcgcaacccaacgcgc	2394
Db	2314	AATCTTATGCAAAATGGCGCATTAACGCTTATACAGTATAGCCACACGCCACCAACCAACGCG	2373
QY	2395	aacctcagcctcgttggccaatgcccaagcaaacattlaatcaagccacatlaaagcgcaac	2454

Db 880 CAGCTGATACGCAAAAGTGTCTACGATGACATTTACAGAGCCGATACACTACCCGTC 939
Qy 961 tctacgaacacatcaaaatgggaatatttcttaacagacaataaataaggcagaga 1020
Db 940 TMTTGTGAACCGGCGAGTAAACGACATTTTCTTACATCCAAACAAACGCTACGGGT 999
Qy 1021 aaatcagtcgaacaaataacacattctctacattatagattaaacagacagcgtt 1080
Db 1000 ACGGTAAACAAACAAACGAAAGGTTTCATCC---AAAGCTTAAAGTACAGACATC 1056
Qy 1081 caattgttaatgttcttcttaacagagacagagaaacgttcttaacgtcaggt 1140
Db 1057 CGAGTGTGTAACAACTCTTGAATGAATGAAACGATGAAGAACCACTTT---ACGGCGAGGG 1113
Qy 1141 ggggtcacacgttatccagaccagactgaataatggagaataatcttccattatgaca 1200
Db 1114 GGTGTTAATGATACGCTCCAGGTTAAACAGGTGAACAACTTTCTTTATCGATTTAC 1173
Qy 1201 ggaagaagttaattgtactactaccagacaatacaacaaagcgcgaggttcttct 1260
Db 1174 GGCACAGCGCAACTCATCTTATCAACACATCAACCAAGCGCGGGGCTTTGTATTTT 1233
Qy 1261 gaaggttaatttaacgttctcgtccttaaaacaaacgaacgttgaagcgcggtctcat 1320
Db 1234 GAAGGTGATTTTACGGTCTGCTCTGAAACACAGCAACGTCGCAAGCGCGGGCTTCAT 1293
Qy 1321 atcagtgatgcaggtaccgttacttgaagaagttaacgsggtgacaaacagcgtctgc 1380
Db 1294 ATCAGTGAACAGATGATACGCTTACTTGGAAAGTAAACGCGCTGGCAACGCGCTGTCC 1353
Qy 1381 aaaaatgcgaacgaacgctcgtcgttcaagcgaaggggaaacaaacgaacgtcgtcagc 1440
Db 1354 AAAATCGCGCAAGGCGACGCTGCAAGCTTCAAGCCAAAGGGGAAACCAAGCTGATCAGC 1413
Qy 1441 gtgagcagcgttaaaatcatcttgaatcagcagcgagcagatcaaaagcaaaacagcc 1500
Db 1414 GTGGGCGACGCTACATGATCTTTGGATCAGCAGCAGCAGCATTAAGCAAAAACAAAGCC 1473
Qy 1501 ttaagtgaatcagcgttgcagcagcagcagcagcagcagcagcagcagcagcagcag 1560
Db 1474 TTTAGTGAATTCGGCTTGATGATCAGCGGAGGATGATGATGATGATGATGATGATGAT 1533
Qy 1561 ttaacccgcaacacactctctcgtcgttgcgagcagcagcagcagcagcagcagcagc 1620
Db 1534 TTCAACCCGCAACACTCTATTTTCGGCTTTCGGCGGAGGATTTGAATTAACGGGCAAT 1593
Qy 1621 tgcgttgcgttccacgcagcagcagcagcagcagcagcagcagcagcagcagcag 1680
Db 1594 TCCCTTTCGTTCCACGCTATTCAAATACCGATGAAGGGCGATGATGATGATGATGAT 1653
Qy 1681 caagcaagaatccacacgcttaccatctacagcagcagcagcagcagcagcagcagc 1737
Db 1654 GCGCAACAAACATCCACGCTTACCATTTACAGGGAATGAAGTATTCACACACGATGGT 1713
Qy 1738 ---aataacacaacttgcagcagcagcagcagcagcagcagcagcagcagcagcag 1794
Db 1714 AAGAAATATCATGATTAATTAATTAACAAAGAAATTCCTTCAACAGCTTGTGTTGCGAG 1773
Qy 1795 aagatgaacaaacaaacgaacgagcagcagcagcagcagcagcagcagcagcagcagc 1854
Db 1774 AAGATGACGACCAAAACGAGGCGGCTTAACCTTGTATACAGCGCGCGCGAGAGAGAG 1833
Qy 1855 cgaacttactcgttccgagcagcagcagcagcagcagcagcagcagcagcagcagcag 1914
Db 1834 CGACACCGCTGCTTCCCGCGGAGCAAAATTAACGGCAATCAACGCAACAAACGCGC 1893
Qy 1915 aaactgttttccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1974
Db 1894 AAACGTGTTTTCAGCGGACAGCAGACACCGCTACATATATTAGAAACGGGCTG 1953
Qy 1975 tcaaaaatggaaggtatcccaagaaggaatcgtgtggaacaaagattgcagcagcagc 2034
Db 1954 TCAAAAATGGAAGGTATCCCAAGAGAGAAATCGTGTGGAGCAACGACTGATCNAACCGC 2013

Qy 2035 acatttaagcggaaacttccatattcagggcgagacagcagcagcagcagcagcagcagc 2094
Db 2014 ACGTTTAAAGCGGAAATTTTCATATTCAGGCGGAGCGAGCGAGTATTTCCGCAATGTT 2073
Qy 2095 gccaaagtgaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2154
Db 2074 GCCAAGTGGAGGCGCATTTGACATTTGACATTCAGCCCAAGCAAGCTTTTGTGTCCGA 2133
Qy 2155 ccgcatcaaacgcaacacactcgtacagcgttgcagcagcagcagcagcagcagcagcagc 2214
Db 2134 CCGCATCAAAAGCATATCATCTGTACAGCTTGGATGAGTGCAGCTGATGATGATGATGAT 2193
Qy 2215 gaaaacaccattacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2274
Db 2194 GAANNAANATTTNACGAGCATTAAGTATGATGATGATGATGATGATGATGATGATGAT 2253
Qy 2275 aatgtcagccttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2334
Db 2254 ANTGTAGCTNNCCNATNACGTTNTTNAANCTCNCNGGCTNCCNACTNANAGGC 2313
Qy 2335 aatctagtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2394
Db 2314 AATCTTAGCAAAATGCGCATACGCTTATACAGCAGCCACCAAGCAACCAACCAAGCGC 2373
Qy 2395 aacctcagccttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 2454
Db 2374 AACCTTAGCTGCTGCGCATGACGATCCCAAGCAATTTAATCAAGCCACATTAACGCGAAC 2433
Qy 2455 acatcgttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2514
Db 2434 NCATCGGNTTCGCGCATATCTTCAATTAATCAACCAACCAAGCGCGCAT 2493
Qy 2515 ctgagccttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2574
Db 2494 CTGACGCTTTCGACAGCAGCTTAAGGCAAGGATGATGATGATGATGATGATGATGATGAT 2553
Qy 2575 tccctagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2634
Db 2554 TCCCTAGCGGATTAAGCATATTTCAATTTGAACAGCGCGCTTACGGAACATCTCAGC 2613
Qy 2635 ggcggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2694
Db 2614 GCGACGACAGGACAGCATTTACATTTAAGAGACAGAGATGAGACCTGCGTCAGGACAGC 2673
Qy 2695 gaattagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2754
Db 2674 GAATTAAGCAATTTAAACCTTGACAAACGCGCATTTACACTCAATTCGCTTATGCGCAC 2733
Qy 2755 gatcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2814
Db 2734 GATGCTGACAGGCGGCAACCGGCAAGGATGATGATGATGATGATGATGATGATGATGAT 2793
Qy 2815 tccctattatccgttaccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2874
Db 2794 TCCCTATTATTCGTTTACACCGGCAACTTCGATGATGATGATGATGATGATGATGATGAT 2853
Qy 2875 aacggaacattgaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2934
Db 2854 AACGGAATTTAAGCAGTGAAGCAATTCGCTTATGATGATGATGATGATGATGATGATGAT 2913
Qy 2935 agcgcgaacattgaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2994
Db 2914 ACGGCAAAATTAAGCTGCGGAAAGTTCCGAAGNACTTACCTTGTGCGCTTAACAT 2973
Qy 2995 accggaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3054
Db 2974 ACCGGCAAGCAACCGGTAAGCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 3033
Qy 3055 ccgctgtccgaaatcttaatttcaacctgcagcagcagcagcagcagcagcagcagcagcagcagc 3114
Db 3034 CCGCTGTGGAACCACTTAATTTACCTGCAAAACGAAACAGATGATGATGATGATGATGATGAT 3093

TITLE Smith, H.O. and Venter, J.C.
Whole-genome random sequencing and assembly of Haemophilus influenzae Rd
JOURNAL science 269 (5223), 496-512 (1995)
MEDLINE 95350630
REFERENCE 2 (bases 1 to 13818)
AUTHORS Tatusov, R.L., Mushegian, A.R., Bork, P., Brown, N.P., Hayes, W.S., Bordovsky, M., Rudd, K.E. and Koonin, E.V.
TITLE Metabolism and evolution of Haemophilus influenzae deduced from a whole-genome comparison with Escherichia coli
JOURNAL Curr. Biol. 6 (3), 279-291 (1996)
MEDLINE 96398784
REFERENCE 3 (bases 1 to 13818)
AUTHORS White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.
TITLE Direct Submission
JOURNAL Submitted (25-Jul-1995) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA
REFERENCE 4 (bases 1 to 13818)
AUTHORS White, O., Clayton, R.A., Kerlavage, A.R. and Fleischmann, R.D.
TITLE Direct Submission
JOURNAL Submitted (27-Sep-1997) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA
REMARK The H. influenzae sequence has been updated by R. Fleischmann. New database matches have been assigned, product names have been improved, and a number of frame shifts have been corrected. We gratefully acknowledge the work of Tatusov et. al. We have incorporated their annotation into the /notes fields of the corresponding H. influenzae genes
5 (bases 1 to 13818)
AUTHORS White, O., Clayton, R.A., Kerlavage, A.R., Fleischmann, R.D., Peterson, J., Hickey, E., Dodson, R. and Gwinn, M.
TITLE Direct Submission
JOURNAL Submitted (28-May-1998) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA
REMARK The whole genome was shifted by 588 nucleotides for a new start
COMMENT On Sep 30 1996 this sequence version replaced gi:1220911.
FEATURES
source
Location/Qualifiers
1..13818
/organism="Haemophilus influenzae Rd"
/db_xref="taxon:71421"
complement(134..343)
/gene="HI0235"
complement(134..343)
/note="similar to GB:129458 SP:P36675 PID:473421 percent identity: 73.81; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC21905.1"
/db_xref="GI:1573201"
/translation="MRKKSASVENETVYLTHTGVYKDNVAVALLGDKLFRQVREKKR KGGSYORAKHPGKIFPEKPDYKFF"
complement(411..761)
/gene="HI0236"
complement(411..761)
/note="similar to PID:1805555 percent identity: 45.74; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="arsenate reductase, putative"
/protein_id="AAC21906.1"
/db_xref="GI:1573202"
/translation="MSVITTHNPHCSKRETLALDKNGIPIITELYIQKQSYNELO SIATKGLIDYDROMKRTDELYKSLINDLIDLSQAEIYFKAISHSALIERIVINGDK AKIKRPETVLEIL"
840..1888
/gene="HI0237"
/note="This region contains an authentic frame shift and is not the result of a sequencing artifact: similar to GB:U00096 SP:P77406 PID:1621505 PID:1788838 percent

identity: 55.75; identified by sequence similarity; putative"
/pseudo
complement(2003..2980)
/gene="HI0239"
complement(2003..2980)
/gene="HI0239"
/note="similar to SP:P19674 GB:X56175 PID:581231 GB:U00096 PID:1773093 percent identity: 48.00; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="protein-export membrane protein (secf)"
/protein_id="AAC21907.1"
/db_xref="GI:1573204"
/translation="MMKLFTRKDGCHTIREINGIKLFFPIITPKYKAKGIIISALLM VTSLEFIIITKGFNGLDTGVVDFTHFSQANLEQIRSKLHENGISPIVQYTSVO DVMIRLPASNDSTIGEHVSMLOVDKIDQIRSEFVNGEELAGAVATLATL AMVLIVYSREMRGSGIASLAHDVITIGVFSALQIEDLFFVAIISVVGSSIN DSIVFDFRVRNFRKIRRLDTIDIDISLTQTSRTIITSVTLLVVMALFFFGPSI HNFSLALVGGIGFCTSSIFVALIAVDVGLRREHMPKVDKIDELP"
complement(2988..4838)
/gene="HI0240"
complement(2988..4838)
/note="similar to SP:P19673 GB:S68715 GB:X56175 PID:581230 PID:545175 percent identity: 59.74; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="protein-export membrane protein (secp)"
/protein_id="AAC21908.1"
/db_xref="GI:1573205"
/translation="MLNRYPMKRMKLVPIFVAGILYSLPIYGEPAVOISGRQGE ANTSVLCQVDVLTNNLPKSTYLENGSLARTTDDDLAKDKIAEPLGNNTTAL ENLPATPAMLSMFGANPMKMGDLRGVFLMEVDNATLVROEDQDSLNLSTVA LEIOYTAIKTNEHFGTLITLVANSORAKARIIROLPTDITPEPDSINLISLTA LNEARDAIEONLTILRRVAELGVAIVARORGAERIVTELGVODTARAKETLIGT ATLEERIVONVNTADAISSRMPLPADSEVKYDRQCHPVALPRKAVLGGEHIIINSSGID QHSSTPOVSTRLDSEGEINISQTKKRYKKPMATLYEYKDNKKDKNGTIIIEKHE VINAVATIOGRGNSNFOITGVDSIAEHNLTSLKSGALAPIDIVERTIGPSLGN VEOGINSIMGVAVIAFMLEFYKMGVIVASPLVNIIVLVGMSITLPGATSMGCI AGIVLTGMSVDANVLIEERIKERIRNGRSIOAINGVNGARTSPDANVTITVVAL ILVAVGCPIDGPAITLISGVAISMFAITGTRALVALYGGQKLLI"
complement(4908..5201)
/gene="HI0241"
complement(4908..5201)
/note="similar to PID:599589 percent identity: 62.37; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC21909.1"
/db_xref="GI:1573206"
/translation="MEASPSMSTLFTFVIGLIFPYPIYTPQAKRKKKKLMSEIAK GTEVLVAGVIGVIGTKTYTESGSDVIALNDTFTITINRYIVSLPGSLKS"
complement(5309..5530)
/gene="HI0242"
complement(5309..5530)
/note="similar to GB:L77117 SP:058397 PID:1591652 percent identity: 35.71; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAC21910.1"
/db_xref="GI:1573207"
/translation="MKYOLNLTALRCPILPLSAKKALKNDKDELMLILNLSAVEN FSIYFAENSVAVLYBOYVASEKEPVLIAKK"
complement(5527..6045)
/gene="HI0243"

Db 504 ATTGTTACGAGAGCGGCTCCATGTAFTAGCTTCGAAATATGAATGCGACTTATTTC 563
QY 558 tgaatttaataataatcaccgatcggttgcgaatcgagacgaacatattgcygctc 617
Db 564 AGGTACAGCAAAATATTCAGAACGTTGCTGATCGCTGAGCGCACTTTGGGGA 623
QY 618 tgaagaagcgaacccaataacccggaagttcaatcatttgcagaagcgaatcttgc 677
Db 624 TGATTCAGAGCAAAAGGC-----GACCAAGTTGGCCGGCTCATATCATTA 665
QY 678 gctcgctcggtgaatcacttgcgaataatgaltggttgcgaagtcgaactgaag 737
Db 666 TCTGACAGCGGCAATACACACAAATCAGCGGTGAGCAGTATGATATTCGATTGGG 725
QY 738 tagcgaataataaacaagccatagtttcttaccacaagagctcatttggcga 797
Db 726 AGCGGATGTTCTGTAAGCGGAGAAATATGTCGATTGACGAGCTCAAGGGGA 785
QY 798 cagtgcacacaaagtctatctatgctgcacaaagcaaaagtggtaattatgggt 857
Db 786 CAGTGTCTCTCCATGTTATTTATGATCTGAAAACAAAATGTTTAAATGCGAT 845
QY 858 atgcgaacagcgaaccccatataggaagaagcaatgcttcgaactgaactgaaga 917
Db 846 ATTACGGGAAGGCAACCTTTTCAAGCAAGAAATGCGTTTCAATGCTTCGCAATC 905
QY 918 ttggtctatgataatcttgcgtgagatacccatcagtaattctacgaacacata 977
Db 906 TTTATTT---TGATGAATTTTTCGAAGAGATTACATACATTCATACCCGAGCTG 962
QY 978 aaatgggaatattcttcttcaacgaataataatgacgagaaatgatgcgaaca 1037
Db 963 TAATGAGGTGACACATTAATGATGATATGATGTCAGGGGTAT-----AACT 1014
QY 1038 taacacttctcctcactatagattaaacacgaacgcttcaattgttcaatgttc 1097
Db 1015 CAGAAATTCAGGAATACCATTCAGAAATTAATTAAC-----TTAGCAAAATGAG 1064
QY 1098 ttatccgaagacgaagaagacgttctatcatgctgca---gltgggtcaacagta 1154
Db 1065 TTTTACCTTTGAAAGAGAGATTAAGTTATCTATATCTTAATATGAGCGACTTAATTA 1124
QY 1155 tcgacacgaactgaataatggaagaataatcttctatcttgcgaagaagaagtgaaat 1214
Db 1125 TTCTCCACGTTTAACAAATGAGAGAAAGCTATATTTATGATCAAAACAAAGATCAT 1184
QY 1215 gatacttaccgaacacatacaacgaagcgcggttgcgtatcttgaaggtatctac 1274
Db 1185 AATCTTCGATTCGATTAACCAAGGCGGGGTGCTCTTTATTTGAGGGTAATTTTAC 1244
QY 1275 ggtctgcctaaacgaacgaagctggaagcgcggtcttcatatcagtgatgcag 1334
Db 1245 AGTATCTCCAAATTTTAAACCAACTTGCGCAGAGGCTGSCATATCAATGTAAGTAATG 1304
QY 1335 taccgttacttggaaagttaaagcggttgcgaacgacgctgcgtcaaaatgccaag 1394
Db 1305 CACCGTTACTTGAAGATTAATAGCGCTGGAACATGTCACATTTCTAAATATGGTAAG 1364
QY 1395 cagcgtcgtggttcaagcgaagaaggggaacgaagctgcgtcagcggtggaagcgtta 1454
Db 1365 AATCATTCAGCTTCAGCAAGGAGGGAATTAAGGTTGATTCAGCGGTAGCGATGTAA 1424
QY 1455 agtcaacttagatcgaagcggaacgaataaagcaaaaacgaagccttaagtaactg 1514
Db 1425 AGTCATTTTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1484
QY 1515 cttgtcgaagcgaggggaagcggttgcgaactgaatgcgaatacagttcaacccgaca 1574
Db 1485 CTTGGTTTACCGGAGAGGAGCTGTTCAATTAACGATGATTAACATTTGATACCGATTA 1544
QY 1575 acctattctcgttcgagcgagcggttgcgaactgaatggaagcgaactgcgttcgca 1634
Db 1545 ATTATTTTCGGGTTTCGATGAGGTGCTGCTTAATCTTAACGGGCAATTCATTAACCTTTAA 1604

QY 1635 ccgattcaaaataacgatgaagggcgatgatgtgtaacacacatcaagaacaagatc 1694
Db 1605 AGGTATCCAAATATACGAGCGGGGCAATATGTTGACCATTAATACATCAAGCCG 1664
QY 1695 caccgttaccattacaggaataaa-----gatattactacaacggccaataacaaca 1748
Db 1665 TAATGTACTATTTACTGGGAACGAAACCATTTGTCATTAATGAAATTAATTAATA 1724
QY 1749 ctgtgatacaaaaagaatctgcctacaacggttgcgttgcgaagaagatgaacaa 1808
Db 1725 ACTGATTACGAAAGAAATGCTTCAACGAGTGGTTGGCGCAACAGATTAATAATA 1784
QY 1809 aacgaagcggtgcctcattcgaattaccacgaaggaagcgatgccttactgc 1868
Db 1785 ACACAAATGGGATTAACCTTATTTTAACCAACCAACCAACCAACCAACCAACCAAC 1844
QY 1869 ttccgcggaacaaatttaaaacggaatatacgaacaaacgaacgcaactgtttcag 1928
Db 1845 TTTAGGTGCTTAATTAATTAAGCGCATATTTACCCAAACAAAGGTAACTATTTTACG 1904
QY 1929 cggcgaacccgaacccgcctacatcatttagaagcggttgcgaagaatggaag 1988
Db 1905 CGGTAGACCGACACCGCCTCATCATTTTAATTAACGTTTGCACAAATGGAAG 1964
QY 1989 tatcccaaggaagaatcgttgcgaacgaatggtatgcgcgcgaactttaagcgga 2048
Db 1965 TATACCAAGAGCAAAATGTTGAGATCACGATTTGATCAACCGTTCATTTAAAGCTGA 2024
QY 2049 aaacttccatatacgaagcggaacgaagcgttgcctccgaatgttgcgaagtggaag 2108
Db 2025 AAATTCCTCAATTAAGCGGAGTGGGATGCTTTCGCAATGTTCTTCAATTTGAGG 2084
QY 2109 cgatttgattgaagaacacgccaagcaagtttgcgttgcgaacggaatcaagca 2168
Db 2085 AAATTTGACATCAGCAATTAATGCAAAATGCCACATTTGGTGTGTCGCAATACAAA 2144
QY 2169 cacaactgtacacgttcgacgttgcgaagcgttgcgaagttgtaacgaaaaacacatc 2228
Db 2145 TACCATTTTGCAGCGGTTCAGATTGACAGGATTAACCACTTGTCAAAAAGTGAATTAAC 2204
QY 2229 cgaacataaagtgttgcctatgcgaagcaagcgaacgaacgaatgcgaactgc 2288
Db 2205 CGATTAACAAATTTATTTATTTATTAACCAAAACCAATCAATGCGCTTATTAATTAAC 2264
QY 2289 cgaacacgtcatttaactcgaaggaactgcgaacacgaacgaacacgttgcgaag 2348
Db 2265 TGATTAATGCAACGCGGATTTAAAGGTTTAAGCAAACTTAATGCAATGCTACTTAAC 2324
QY 2349 cgaagacacgacactatacgttgcgaagcgaacgaacgaacgaacgaacgtcgt 2408
Db 2325 AAATTCACAGCA----- 2336
QY 2409 gggcaatggcccaagaacatttaacgaacacataaagcaaacacatcgttcgga 2468
Db 2337 ----- 2336
QY 2469 caatgctcatttaactcaagcaacgaacgcgtacaaacggaagtcgcgttcgga 2528
Db 2337 -----ATTTCATTAAGCAACAAATGCGCACCAATATGAGCAATTTGACTTTCGA 2387
QY 2529 caagcgaaggaacgaactgaagcacttcgcgaactcaaggaaggtcctcctagcgata 2588
Db 2388 CATTTCAACTGCAACGGTGGATTAATGCAAACTGAAACGTAATGATGATTAACGATTC 2447
QY 2589 ggcagatccacttgcgaacacgcttaccggaanaatcagcgcggaagatc 2648
Db 2448 AGCTCAATTTCTTTAAATAACACCCATTTTTCGACCAATTTCAAGGAGCAAAAGGCAC 2507
QY 2649 ggcattcaacttaaaagacagcggaatgacgttgcgttcggaacggaatgaagcaatt 2708
Db 2508 AACAGTGAAGTTGGAATAATGAGACTTGGAACAATGCCATGAGATCAATTCAGAAATTT 2567

QY 2709 aaaccltgaacagccacccatcaactccgctatcgacacgagtcgagcg 2768
 DB 2568 AACCTAAATTAACGATCGATCGATCAAGTTAAATTCAGCTTATTC----- 2610
 QY 2769 gcaaacgagcagtcgagcagtcgagcgccgctccgagcttccctatccgt 2828
 DB 2611 -----GCTAGCTCAACAAATACGCCACG-----TCGCCGTTCAATTAGACGGA 2654
 QY 2829 tagcgccgcaacttcgagcaaatccgcttcaacacgagcgagtcgaacattga 2888
 DB 2655 AACAAACGCAATCGACATGCAATGCTTCAACATGACATGAAATGATGAG 2714
 QY 2889 cgtcagaggaacatccgcttcatcggaactccgctacacgagcggaattga 2948
 DB 2715 TGGCAAGGCAATTCCTTACTTACTTATTTGGCTTAAAGGATTAATTA 2774
 QY 2949 gctgaggaagatccggaagcacttaacacttcgctgaacaaatccggaacacc 3008
 DB 2775 ATTATTCATGACGCTGAGGCGCATTAATATATCTGTTCCGAACACAGCGAAAGAAC 2834
 QY 3009 cgtagatcgtgagcaattgscgtagtgaagaaagaaacacacacgcttccgaaa 3068
 DB 2835 CGAAACCTTGTGACATTAATTTGTTGTAACCAAGATTAATCAACGCTTATCAATTA 2894
 QY 3069 tcttaattcaacctgcaaaacgaacacgctcgaagtcgagcgagtcgagtcacat 3128
 DB 2895 GCTCAAAATTTACTTTAGAAATTAACACGCTGATCGAGCTGATTAATTAATTAAT 2954
 QY 3129 ccggaagagcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 3188
 DB 2955 GAAGAAATGATGCGAATTTCCGCTTGCATTAACCATTAAGAGAGAGATTCACACATGA 3014
 QY 3189 actgagcaagcgaggaagaaacgagcgcttgaagcaagaaacgagtcgagtcgagtc 3248
 DB 3015 TTTAGTAAAGCAGAGCAACAGCAACATTAAGAAACCAAGATTAAGAAACCAAGTC 3074
 QY 3249 caaacacagcgaggaagaaacgagcgagtcgagtcgagtcgagtcgagtcgagtc 3308
 DB 3075 TAAACACAAACAGATGAGCAAAAGTGCCTGTCAGAAAGAGAGAGAGAGAGAGAG 3134
 QY 3309 caatgcaacggaagaaagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 3368
 DB 3135 TGATGATCCCTGCTGATCAAAAGCTGTTAAAGCATTAAGAAAGCAACAGTGAAGTC 3194
 QY 3369 tgcgagcatatgagcgaggaagaaacgagtcgagtcgagtcgagtcgagtcgagtc 3428
 DB 3195 TGTGAAACAAACAAAGTAAAGCAAAACAAAGTAAAGCAAAAGTAAAGCAAAAGT 3254
 QY 3429 ctgagcaaacagcgaggaagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 3488
 DB 3255 TTTGATGATCCCTGCTGATCAAAAGCTGTTAAAGCATTAAGAAAGCAACAGTGAAGTC 3314
 QY 3489 ccgagccgagcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 3546
 DB 3315 TGCCCCACAGCAATCGGAAGAAAGTGTAGCTCAAGAAAGGAGGAGGAGGAGGAG 3374
 QY 3547 -----gactgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 3599
 DB 3375 ACAAAAGACGTCGATGAGCGGTTATTCAAATGATGATGATGATGATGATGATGATG 3434
 QY 3600 caaacagcttccgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 3659
 DB 3435 AAATAGTATGCTTCTTCTTCAAGATGATGATGATGATGATGATGATGATGATGATG 3494
 QY 3660 cgcgcttgaacagcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 3719
 DB 3495 TGCCGCTGAGCAAAATTCGACAGGATTAAGCAACGCTTATGATGATGATGATGATG 3554
 QY 3720 ctacgagcaaacagcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 3779
 DB 3555 TTATCAGCAAGAAAGCAATTAAGTAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3614
 QY 3780 cgtcagcatcctgcttccgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 3839

DB 3615 AATTGGGCGAGTTTCTTCGATAGCGCTTCAGATATATACCTTGTGATGAACAGGTTAA 3674
 QY 3840 ctgcgagcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 3899
 DB 3675 TCAGCGGACATTAAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 3734
 QY 3900 cgtcagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 3959
 DB 3735 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3794
 QY 3960 aatcgcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 4019
 DB 3795 AATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3854
 QY 4020 atcgcagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 4079
 DB 3855 ATTGGGATTCAGCCTTATTTGAGATTAATGCTATTTTATTAAGCTGAAGATTAATTA 3914
 QY 4080 atcgaagacgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 4139
 DB 3915 ATCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3974
 QY 4140 taagcagatattcaatcaaacgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 4199
 DB 3975 TCGAGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4034
 QY 4200 gtcctatccgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 4259
 DB 4035 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4094
 QY 4260 gcaagatattcgaagaaacgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 4319
 DB 4095 ACATACATTTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 4154
 QY 4320 cagcgtgctccacagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 4379
 DB 4155 CCAAAATTTCCGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4214
 QY 4380 gggcatcaatlagctacgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtc 4407
 DB 4215 GGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4242

RESULT 14
 LOCUS AR157320 4319 bp DNA linear PAT 17-OCT-2001
 DEFINITION Sequence 1 from patent US 6245337.
 ACCESSION AR157320
 VERSION AR157320.1 GI:16218252
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1. (bases 1 to 4319)
 AUTHORS St. Geme J.W., III and Falkow S.
 TITLE Haemophilus adherence and penetration proteins
 JOURNAL Patent: US 6245337-A 1 12-JUN-2001;
 FEATURES
 Source 1..4319
 BASE COUNT 1497 a 776 c 891 g 1155 t
 ORIGIN

Query Match 29.4% Score 1297.8; DB 6; Length 4319;
 Best local Similarity 59.1%; Pred. No. 1e-293;
 Matches 2571; Conservative 0; Mismatches 1567; Indels 213; Gaps 12;

QY 78 ttcctagcagcatcgcctgctgctgagcatttcgcccaagcgagcgagcgagcaactta 137
 DB 86 TTTTAAACCGCTTGATTAATTAAGGATTAATGCAAGCGTGGCTGCTACACTTA 145

[illegible]

Chr	Start (kb)	End (kb)	Gene	Strand	RefSeq ID	Accession	Length (bp)	GC (%)	GC3 (%)	GC4 (%)	GC5 (%)	GC6 (%)	GC7 (%)	GC8 (%)	GC9 (%)	GC10 (%)	GC11 (%)	GC12 (%)	GC13 (%)	GC14 (%)	GC15 (%)	GC16 (%)	GC17 (%)	GC18 (%)	GC19 (%)	GC20 (%)	GC21 (%)	GC22 (%)	GC23 (%)	GC24 (%)	GC25 (%)	GC26 (%)	GC27 (%)	GC28 (%)	GC29 (%)	GC30 (%)	GC31 (%)	GC32 (%)	GC33 (%)	GC34 (%)	GC35 (%)	GC36 (%)	GC37 (%)	GC38 (%)	GC39 (%)	GC40 (%)	GC41 (%)	GC42 (%)	GC43 (%)	GC44 (%)	GC45 (%)	GC46 (%)	GC47 (%)	GC48 (%)	GC49 (%)	GC50 (%)	GC51 (%)	GC52 (%)	GC53 (%)	GC54 (%)	GC55 (%)	GC56 (%)	GC57 (%)	GC58 (%)	GC59 (%)	GC60 (%)	GC61 (%)	GC62 (%)	GC63 (%)	GC64 (%)	GC65 (%)	GC66 (%)	GC67 (%)	GC68 (%)	GC69 (%)	GC70 (%)	GC71 (%)	GC72 (%)	GC73 (%)	GC74 (%)	GC75 (%)	GC76 (%)	GC77 (%)	GC78 (%)	GC79 (%)	GC80 (%)	GC81 (%)	GC82 (%)	GC83 (%)	GC84 (%)	GC85 (%)	GC86 (%)	GC87 (%)	GC88 (%)	GC89 (%)	GC90 (%)	GC91 (%)	GC92 (%)	GC93 (%)	GC94 (%)	GC95 (%)	GC96 (%)	GC97 (%)	GC98 (%)	GC99 (%)	GC100 (%)	GC101 (%)	GC102 (%)	GC103 (%)	GC104 (%)	GC105 (%)	GC106 (%)	GC107 (%)	GC108 (%)	GC109 (%)	GC110 (%)	GC111 (%)	GC112 (%)	GC113 (%)	GC114 (%)	GC115 (%)	GC116 (%)	GC117 (%)	GC118 (%)	GC119 (%)	GC120 (%)	GC121 (%)	GC122 (%)	GC123 (%)	GC124 (%)	GC125 (%)	GC126 (%)	GC127 (%)	GC128 (%)	GC129 (%)	GC130 (%)	GC131 (%)	GC132 (%)	GC133 (%)	GC134 (%)	GC135 (%)	GC136 (%)	GC137 (%)	GC138 (%)	GC139 (%)	GC140 (%)	GC141 (%)	GC142 (%)	GC143 (%)	GC144 (%)	GC145 (%)	GC146 (%)	GC147 (%)	GC148 (%)	GC149 (%)	GC150 (%)	GC151 (%)	GC152 (%)	GC153 (%)	GC154 (%)	GC155 (%)	GC156 (%)	GC157 (%)	GC158 (%)	GC159 (%)	GC160 (%)	GC161 (%)	GC162 (%)	GC163 (%)	GC164 (%)	GC165 (%)	GC166 (%)	GC167 (%)	GC168 (%)	GC169 (%)	GC170 (%)	GC171 (%)	GC172 (%)	GC173 (%)	GC174 (%)	GC175 (%)	GC176 (%)	GC177 (%)	GC178 (%)	GC179 (%)	GC180 (%)	GC181 (%)	GC182 (%)	GC183 (%)	GC184 (%)	GC185 (%)	GC186 (%)	GC187 (%)	GC188 (%)	GC189 (%)	GC190 (%)	GC191 (%)	GC192 (%)	GC193 (%)	GC194 (%)	GC195 (%)	GC196 (%)	GC197 (%)	GC198 (%)	GC199 (%)	GC200 (%)	GC201 (%)	GC202 (%)	GC203 (%)	GC204 (%)	GC205 (%)	GC206 (%)	GC207 (%)	GC208 (%)	GC209 (%)	GC210 (%)	GC211 (%)	GC212 (%)	GC213 (%)	GC214 (%)	GC215 (%)	GC216 (%)	GC217 (%)	GC218 (%)	GC219 (%)	GC220 (%)	GC221 (%)	GC222 (%)	GC223 (%)	GC224 (%)	GC225 (%)	GC226 (%)	GC227 (%)	GC228 (%)	GC229 (%)	GC230 (%)	GC231 (%)	GC232 (%)	GC233 (%)	GC234 (%)	GC235 (%)	GC236 (%)	GC237 (%)	GC238 (%)	GC239 (%)	GC240 (%)	GC241 (%)	GC242 (%)	GC243 (%)	GC244 (%)	GC245 (%)	GC246 (%)	GC247 (%)	GC248 (%)	GC249 (%)	GC250 (%)	GC251 (%)	GC252 (%)	GC253 (%)	GC254 (%)	GC255 (%)	GC256 (%)	GC257 (%)	GC258 (%)	GC259 (%)	GC260 (%)	GC261 (%)	GC262 (%)	GC263 (%)	GC264 (%)	GC265 (%)	GC266 (%)	GC267 (%)	GC268 (%)	GC269 (%)	GC270 (%)	GC271 (%)	GC272 (%)	GC273 (%)	GC274 (%)	GC275 (%)	GC276 (%)	GC277 (%)	GC278 (%)	GC279 (%)	GC280 (%)	GC281 (%)	GC282 (%)	GC283 (%)	GC284 (%)	GC285 (%)	GC286 (%)	GC287 (%)	GC288 (%)	GC289 (%)	GC290 (%)	GC291 (%)	GC292 (%)	GC293 (%)	GC294 (%)	GC295 (%)	GC296 (%)	GC297 (%)	GC298 (%)	GC299 (%)	GC300 (%)	GC301 (%)	GC302 (%)	GC303 (%)	GC304 (%)	GC305 (%)	GC306 (%)	GC307 (%)	GC308 (%)	GC309 (%)	GC310 (%)	GC311 (%)	GC312 (%)	GC313 (%)	GC314 (%)	GC315 (%)	GC316 (%)	GC317 (%)	GC318 (%)	GC319 (%)	GC320 (%)	GC321 (%)	GC322 (%)	GC323 (%)	GC324 (%)	GC325 (%)	GC326 (%)	GC327 (%)	GC328 (%)	GC329 (%)	GC330 (%)	GC331 (%)	GC332 (%)	GC333 (%)	GC334 (%)	GC335 (%)	GC336 (%)	GC337 (%)	GC338 (%)	GC339 (%)	GC340 (%)	GC341 (%)	GC342 (%)	GC343 (%)	GC344 (%)	GC345 (%)	GC346 (%)	GC347 (%)	GC348 (%)	GC349 (%)	GC350 (%)	GC351 (%)	GC352 (%)	GC353 (%)	GC354 (%)	GC355 (%)	GC356 (%)	GC357 (%)	GC358 (%)	GC359 (%)	GC360 (%)	GC361 (%)	GC362 (%)	GC363 (%)	GC364 (%)	GC365 (%)	GC366 (%)	GC367 (%)	GC368 (%)	GC369 (%)	GC370 (%)	GC371 (%)	GC372 (%)	GC373 (%)	GC374 (%)	GC375 (%)	GC376 (%)	GC377 (%)	GC378 (%)	GC379 (%)	GC380 (%)	GC381 (%)	GC382 (%)	GC383 (%)	GC384 (%)	GC385 (%)	GC386 (%)	GC387 (%)	GC388 (%)	GC389 (%)	GC390 (%)	GC391 (%)	GC392 (%)	GC393 (%)	GC394 (%)	GC395 (%)	GC396 (%)	GC397 (%)	GC398 (%)	GC399 (%)	GC400 (%)	GC401 (%)	GC402 (%)	GC403 (%)	GC404 (%)	GC405 (%)	GC406 (%)	GC407 (%)	GC408 (%)	GC409 (%)	GC410 (%)	GC411 (%)	GC412 (%)	GC413 (%)	GC414 (%)	GC
-----	------------	----------	------	--------	-----------	-----------	-------------	--------	---------	---------	---------	---------	---------	---------	---------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	----

DEFINITION Neisseria lactamica Iga protease homologue gene, isolate NL3327.
ACCESSION AJ001740
VERSION AJ001740.1 GI:3152399
KEYWORDS Iga protease.
SOURCE Neisseria lactamica.
ORGANISM Neisseria lactamica
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.

REFERENCE 1 (bases 1 to 578)
AUTHORS Sayers, J.R.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1997) Sayers J.R., Division of Molecular
Medicine, University of Sheffield, Royal Hallamshire Hospital,
Sheffield, S. Yorks, S10 2JF, ENGLAND
2 (bases 1 to 578)
REFERENCE Sayers, J.R., Witowski, S. and Read, R.
AUTHORS
TITLE Strains of Neisseria lactamica carry an Iga1 protease gene
homologue and some express Iga1 protease activity
JOURNAL Unpublished

FEATURES
source
Location/Qualifiers
1..578
/organism="Neisseria lactamica"
/isolate="NL3327"
/db_xref="taxon:486"
<1..>576
/codon_start=1
/transl_table=11
/product="Iga1 protease"
/protein_id="CA004965.1"
/db_xref="GI:3152400"
/db_xref="SPTREMBL:O69746"
/translation="GKFAVGANDIEVNNKGLVSKSMTRKAPMIDSVSRNGVALV
GQIVIVAHNGGVNNDVFGAGSNPDHFRSYQIVKRNRYKPDNSHPYNGDHPRL
HKFVTDAPVEMTGDMRGNTYSDEKERYPERVIRGSHHWRYDDDKHGLSYSGAMLI
GNTHMOCGNGNGVVSLSGDVHNDYGMPI." 88..95
repeat_region
repeat_region 139..148
BASE COUNT 163 a 125 c 156 g 134 t
ORIGIN

Query Match 7.5%; Score 328.4; DB 1; Length 578;
Best Local Similarity 84.7%; Pred. No. 3.5e-66;
Matches 381; Conservative 0; Mismatches 66; Indels 3; Gaps 1;

QY 184 ggcaagtttgcagtcggcgcaagaatattgagtttacaacaaaaggaggttgctc 243
DB 1 GGCACGTTTCAGTCGGGCGCAAGAAATGAGTTTACACAAAAAGGAGGAGTTGTC 60

QY 244 ggcaaatcgatcgagaaagcccgatgatgttcttctgtgtatcgcgtacagcgctg 303
DB 61 GGCACATTCATGACAAAAAGCCCGCATGATGATTTTCGCTGTGTCGCTAACGCGCTG 120

QY 304 ggcgcatgscggcgatcaatatattgtgaggttgacataacggcgctatacaat 363
DB 121 GCGCATTGCTGGGCGCATATATTTGAGGTGCGCATACGCGGCTATTAACAAT 180

QY 364 gtgattttgtgctgagaggaagcaatcccgatcagaccgcttcttccaattgtg 423
DB 181 GTTGATTTTGTGCTCAAGAGGAACCAATCCGATCACACGCGCTTTCTTACCAATTGTG 240

QY 424 aaaaagaataatataaagagagactaagcgcatccttaattgscggagataatag 483
DB 241 AAAAGAAATTAATAAACCCTG---ACAAATTCACACCCTTACACGCGATTAACCATAG 297

QY 484 ccgctgttgcacaaatttgcacagatgcagaacctgttgagattgacagttatgat 543
DB 298 CCGCGCTCTGCGATTAAGTTTCGTAACCGATGCGAAGCTGTGCAATGACGGGTGACATGAG 357

QY 544 gggctgaataagcgtatgaataaaccctgcatcgtgttgaaatcgagagcgaca 603
DB 358 GGGAAATACCTATTCCGATTAAGAAATAATCCGAGGAGTGTCCGATCGGCTAGGACAC 417

QY 604 caatatggcgtctgcatgaagcgaaccc 633
DB 418 CACTATTGGCGTTATGATGATGACAAACAC 447

RESULT 16
LOCUS NLIGA3354 584 bp DNA linear BCT 21-MAY-1998
DEFINITION Neisseria lactamica Iga protease homologue gene, isolate NL3354.
ACCESSION AJ001739
VERSION AJ001739.1 GI:3152397
KEYWORDS Iga1 protease.
SOURCE Neisseria lactamica.
ORGANISM Neisseria lactamica
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.

REFERENCE 1 (bases 1 to 584)
AUTHORS Sayers, J.R.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1997) Sayers J.R., Division of Molecular
Medicine, University of Sheffield, Royal Hallamshire Hospital,
Sheffield, S. Yorks, S10 2JF, ENGLAND
2 (bases 1 to 584)
REFERENCE Sayers, J.R., Witowski, S. and Read, R.
AUTHORS
TITLE Strains of Neisseria lactamica carry an Iga1 protease gene
homologue and some express Iga1 protease activity
JOURNAL Unpublished

FEATURES
source
Location/Qualifiers
1..584
/organism="Neisseria lactamica"
/isolate="NL3354"
/db_xref="taxon:486"
<1..>582
/codon_start=1
/transl_table=11
/product="Iga1 protease"
/protein_id="CA004964.1"
/db_xref="GI:3152398"
/db_xref="SPTREMBL:O69747"
/translation="FAVGANDIEVNNKGLVSKSMTRKAPMIDSVSRNGVALV
GQIVIVAHNGGVNNDVFGAGSNPDHFRSYQIVKRNRYKPDNSHPYNGDHPRLK
FVTDAPVEMTGDMRGNTYSDEKERYPERVIRGSHHWRYDDDKHGLSYSGAMLI
GNTHMOCGNGNGVVSLSGDVHNDYGMPIAGA." 82..88
repeat_region
repeat_region 134..142
BASE COUNT 163 a 127 c 159 g 135 t
ORIGIN

Query Match 7.3%; Score 321.6; DB 1; Length 584;
Best Local Similarity 84.8%; Pred. No. 1.4e-64;
Matches 373; Conservative 0; Mismatches 64; Indels 3; Gaps 1;

QY 190 ttgcaagtcggcgagaaatattgagtttacaacaaaaggaggttgctgcaaa 249
DB 1 TTTCAGTGGGCGCGCAAGAAATGAGTTTACACAAAAAGGAGGAGTTGTCGCGCAA 60

QY 250 tcatatgaacaaagcccgatgatgttcttctgtgtatcgcgtacagcgctgagcgca 309
DB 61 TCAATGCAAAAAGCCCGCATGATGATTTTCGCTGTGTCGCTAACGCGCTGCGCGCA 120

QY 310 ttggcgggcgatcaatatattgtgagcgtgcacataacggcgctatacaattgat 369
DB 121 TTGCTGGGCGCATATATTTGTGAGCGGCGACATACGCGGCTATTAACAATGTTGAT 180

QY 370 ttgtgtcggaggaagcaatcccgatcagaccgcttcttccaattgtgaaaga 429
DB 181 TTGTGTGCGAAGAGGAACCAATCCGATGACGACCGCTTTCTTACCAATTGTGAAGA 240

QY 430 aataatataaagcgtatgaataaaccctgcatcgtgttgaaatcgagagcgcat 489
DB 430 aataatataaagcgtatgaataaaccctgcatcgtgttgaaatcgagagcgcat 489

Db 241 AATACTACAAACCTG---ACAATTACACACCTTACACGGCGATTACCATATGCCGCC 297
Qy 490 ttgcacaaatttgcacaaatgcagaaactgttgagatgacccagttatgagtggtg 549
Db 298 CCGCATTAAGTTCGTAACCGATCGAAGCTGTGGAATATGAGGAGGGAAT 357
Qy 550 aataagccgatttaataataaccctgacgtgttcgaatcgagagcagagcaaatat 609
Db 358 ACCTATTCGATTAAGAAAATAATATCCGAGCGCTGCCGATGCGCAGCAGCAACTAT 417
Qy 610 tggcggtcgtgagagagca 629
Db 418 TGGCGTTATGATGATGACGA 437

RESULT 17
NLGA823
LOCUS NLGA823 572 bp DNA linear BCT 21-MAY-1998
DEFINITION *Neisseria lactamica* Iga protease homologue gene, isolate NL823.
ACCESSION AJ001737
VERSION AJ001737.1 GI:3152389
KEYWORDS Iga protease.
SOURCE *Neisseria lactamica*.
ORGANISM *Neisseria lactamica*.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 572)
AUTHORS Sayers, J.R.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1997) Sayers J.R., Division of Molecular
Medicine, University of Sheffield, Royal Hallamshire Hospital,
Sheffield, S. Yorks, S10 2JF, ENGLAND
2 (bases 1 to 572)
REFERENCE Sayers, J.R., Vitowski, S. and Read, R.
AUTHORS
TITLE Strains of *Neisseria lactamica* carry an Iga1 protease gene
homologue and some express Iga1 protease activity
JOURNAL Unpublished
FEATURES
source Location/Qualifiers
1..572
/organism="Neisseria lactamica"
/isolate="NL823"
/db_xref="taxon:486"
<1..304
/codon_start=2
/transl_table=11
/product="Iga protease homologue"
/protein_id="CAA04962.1"
/db_xref="GI:3152389"
/db_xref="SPTREMBL:O69748"
/translation="KDIEYNNKKGELVGSMTKAPMIDFSVSRNGVAALVGDQIVS
VAHNGYNNVDGAEGBGNDQHRFSYQIVKRNRYKPDNSHPYNGDYMPLRPLPGLSS"
repeat_region
44..51
/rpt_type=INVERTED
BASE COUNT 68 a 124 c 157 g 133 t
ORIGIN

Query Match 6.6%; Score 290.2; DB 1; Length 572;
Best Local Similarity 82.5%; Pred. No. 3,3e-57;
Matches 359; Conservative 0; Mismatches 68; Indels 8; Gaps 2;

Qy 204 gaaagacatgaggttctacacaaagagaggtgttcgcaaatcgatgacgaagc 263
Db 1 GAAGATATTGAGCTTACACAAAAGGAGGAGTGTGCGCAATGACATGACAAAAGC 60
Qy 264 cccgagatgattcttcgtggtatcgtaacgagcggtgagcgatgagcgagatca 323
Db 61 CCGATGATGATTGATTTCCGTGTGCGCTAACGCGGTGCGCATTTGCTGGCGCATCA 120
Qy 324 atatatgtgagcggtgacataacgagcgctataacaatgttgatttgggtgcgaggg 383
Db 121 ATATATTGTGAGCGGTGCGCATACGCGGCGCTATACCAATGTTGATTTGTGCGGAGG 180

Qy 384 aagcaatcccgatcagacgcgtcttcttaccacaaattgtgaaagaaataataaagc 443
Db 181 AGGCAATCCCGATGACGACCGCTTTTCTTCCAAATTTGTAAGAAATTAACRCAAAAC 240
Qy 444 aggaactaacgacatctatgagcgataltcalatgacgagcgttgcacaa 498
Db 241 TG---ACAATTACACCTTACAAACGCGATTAACATATGCTGCTGCTGCAATAG 297
Qy 499 ttgtcagatgacgaagcgtttgagatgacccagttatgagtgatgagtgagaaagc 558
Db 298 TTGCTAACCGATGACGACCTGTGGAATGACGGGTACATGAGGGGAATACCTATTCC 357
Qy 559 gattaaataataaccctgacgtgttcgaatcgagagcagacaaatattgagcggtc 618
Db 358 GATTAAGAAAATAATATCCGAGCGGTGCGCATGCGCTCAGCAGACACCACTATTGCGGTAT 417
Qy 619 gatgaagacgaagcc 633
Db 418 GATGATGACAAACAC 432

RESULT 18
NLGA3293
LOCUS NLGA3293 572 bp DNA linear BCT 21-MAY-1998
DEFINITION *Neisseria lactamica* Iga protease homologue gene, isolate NL3293.
ACCESSION AJ001738
VERSION AJ001738.1 GI:3152385
KEYWORDS Iga1 protease.
SOURCE *Neisseria lactamica*.
ORGANISM *Neisseria lactamica*.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 572)
AUTHORS Sayers, J.R.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1997) Sayers J.R., Division of Molecular
Medicine, University of Sheffield, Royal Hallamshire Hospital,
Sheffield, S. Yorks, S10 2JF, ENGLAND
2 (bases 1 to 572)
REFERENCE Sayers, J.R., Vitowski, S. and Read, R.
AUTHORS
TITLE Strains of *Neisseria lactamica* carry an Iga1 protease gene
homologue and some express Iga1 protease activity
JOURNAL Unpublished
FEATURES
source Location/Qualifiers
1..572
/organism="Neisseria lactamica"
/isolate="NL3293"
/db_xref="taxon:486"
<1..570
/codon_start=1
/transl_table=11
/product="Iga1 protease"
/protein_id="CAA04963.1"
/db_xref="GI:3152386"
/db_xref="SPTREMBL:O69745"
/translation="KDIEYNNKKGELVGSMTKAPMIDFSVSRNGVAALVGDQIVS
VAHNGYNNVDGAEGBGNDQHRFSYQIVKRNRYKPDNSHPYNGDYMPLRPLPGLSS"
repeat_region
67..74
/rpt_type=INVERTED
BASE COUNT 118 a 119 c 154 g 130 t
ORIGIN

Query Match 6.2%; Score 275.4; DB 1; Length 572;
Best Local Similarity 77.0%; Pred. No. 9,9e-54;
Matches 349; Conservative 0; Mismatches 101; Indels 3; Gaps 1;

Qy 205 aaagatatgaggttctacacaaagagaggtgttcgcaaatcgatgacgaagc 264
Db 121 ATATATTGTGAGCGGTGCGCATACGCGGCGCTATACCAATGTTGATTTGTGCGGAGG 180

the complete genome.
AE002424 AE002098
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
1 (bases 1 to 10485)
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, P.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., Deboy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
Cittone, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H.,
Qin, H., Vamathevan, J., Gill, J., Scariato, V., Maignani, V.,
Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappelli, R. and Venter, J.C.
Complete genome sequence of *Neisseria meningitidis* serogroup B
strain MC58
Science 287 (5459), 1809-1815 (2000)
JOURNAL
MEDLINE
PUBMED
2 (bases 1 to 10485)
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, P.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., Deboy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
Cittone, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H.,
Qin, H., Vamathevan, J., Gill, J., Scariato, V., Maignani, V.,
Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappelli, R. and Venter, J.C.
Direct Submission
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
Location/Qualifiers
1. 10485
/organism="Neisseria meningitidis MC58"
/strain="MC58"
/db_xref="taxon:122586"
/note="serogroup: B"
complement(67..795)
/gene="NMB0696"
complement(67..795)
/gene="NMB0696"
/note="similar to GB:M61017 SP:P27675 PID:142988 percent
identity: 81.40; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="amino acid ABC transporter, ATP-binding protein"
/protein_id="AA041113.1"
/db_xref="GI:7225924"
/translation="MIKFNVHKFKDLHYINGVNLKIKKGVVYVCGSGSKSLI
RTVQLSIESEIWMGVNADPKTDLKIRFEGVFGFNLVPHLTVLEMTTAP
MKVGNABOAEKAMLELREVGALHKKKIFPSLGGCOOQVAIRAGLMEPRVIF
DEPTSLDPRMGVYLKMKDLAESGTMNCVTHKMFARFVADRVIFVKGKGLDE
TPRPFPTNPKRRAKQPLDOVMH"
complement(968..1747)
/gene="NMB0697"
complement(968..1747)
/gene="NMB0697"
/note="similar to GB:D10483 SP:P06992 GB:M11054 GB:M68521
GB:X04711 percent identity: 69.41; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="dimethyladenosine transferase"
/protein_id="AA041114.1"
/db_xref="GI:7225925"
/translation="MKHKARKRFGQNFLODTRIIISDINAVNPROADVIEIGPIGA
AITEPLAKKINRLHVEIDIDYICRLTLPFADKLVIHEDGVLODFNGIAGKKKIVG

gene
CDS
1944..2633
/gene="NMB0698"
/gene="NMB0698"
/note="hypothetical protein; identified by glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA041115.1"
/db_xref="GI:7225926"
/translation="MLFGLPCLAWCKNRLSAVLVPGGLVSNLEIETHPDYL
PKPAVMMAGTFPEPKEDRAMQFHYPNQNDMMRVYGLVFNDAHFORSEKAFDE
KIKAFIHERGISCTVLKAVRQHNADSKLKVETDIAVLAKIPEDNICITGG
KATEILIDGGGIMKPKTGTETVPEFAGRDLTITLPSTSRAPVLSAKKAAYRAV
FEMACLEKOL"
2716..3918
/gene="NMB0699"
/gene="NMB0699"
/note="similar to GB:M58444 SP:P16706 PID:141798
PID:461373 percent identity: 80.71; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="tyrophan synthase, beta subunit"
/protein_id="AA041116.1"
/db_xref="GI:7225927"
/translation="MKNYHADPEKGFEGHGLVSETLIPALDELADAVKARNDE
FWEARHDKHYGRSPVYHARLSLEHGAQILKREDLNTGAKHVNNTIGCAL
AKRMGRKRYATVGTAGOHGAVASATVAAFGMTCOVYMGADIDROMPVFPMKILGAN
VYGVSGSRITLKDAMENAMREKVARVDOTFTYIGTAAGPAPYPPVDPPOCVINAEK
AOMQRAIGQDPVAVACVCGSGNAIGTTPHPIGEEVNYLVCEVAGGCVLPDHAAP
TSGAPIGVHGFRTLMODENGVLTGHSVSGLDYPTICGPHSHIDIKRYETVAK
DEALEAPDLICREFGIIIPALSSHAVAVAKNAKMKKDDVITLVNLSGKDDKIDNTV
ALKGIKIL"
complement(4097..9544)
/gene="NMB0700"
complement(4097..9544)
/gene="NMB0700"
/note="similar to GP:732874 percent identity: 93.40;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="Iga-specific serine endopeptidase"
/protein_id="AA041117.1"
/db_xref="GI:7225928"
/translation="MKTRKFINATLSIFLAVLTPYSEALVBDVDYOIFRDEAF
MGKFPVGVATDLSYKNGKQNTGNALSNVPMIDFVADVNRRTLTVIDPOAVSKHV
KDELSIYHHNHGHDLSNDENRYSVQNOYEPKNNHNGHGLDELPYNNARLNKFPV
TEVAPITAPSSAGGVEITKDKNRFSEFVAVGAGTPEVNSRNIMELSRATRYAAGT
PQDVAVVTSNLNOBELIGDKNRSHSPKLEKLEVSQNALTYAVLGSQSPFLAYOK
OKRWMVFLAVLYMAGYONQSWQENIYKKEFADIKORDNAGTILKNGEHNKTKGT
NHSIGTFLAVLANNEDANNGQVFEENNGTLYLDQNNOGAGLFFPGDGHMKTGIN
DITWLGAGIDVDAGKVVYQVKNPNQDNLTKGKTLINGTGVQGOGLKAGDGVIL
NNOADAKGVQVAFQVIGVSGRTLVLSNNOINDNLYFEGRGRLDANGDLFEH
TENVDEGARIVYVNTGHASTITLTKGSLITDKPTISIIYIONDDDDAGYYVRRKP
ITQGRDLTYKRYTALKSGSVNAPMPENQGTENDMDPDEKKNAMNNKNN
QRTSGSGFEGENGHNGALNLNNGSKRQNRRLTLTGCTLNKISVYOGNYLSSG
RPTPHARDVFNSSAKDNHFSKNNVEFEDWIVRTKATEIYANOSASVSSGNVVS
DITANTATDANKVYNGYNGDEVCVRSYDTGYVTCNNGNSDKALSLNFGATQING
NLNQNALVILGAALMGQIOGQNSVSVLSNHSKHLTGDSQVHNLSDLSHILHNA
SDOSANKYHLTKLNLHSGHFHYLTHLAKMLGKVLKVESASGHYDLHVDKGTGP
NOEGLNLPASSVODRSRLSVSLANNHVDGLARTITKENGITELTPYAEKRRRV
PAPSPATNTASOAKATORDGOAIDAKPONTIYVAPSPANOAEKARQOAAEVOYK
QAEAEKSAELAKOKAAREAPARLATAKAEQSSLEIARHHEKREAELEAKOK
VEAEKQALAVARKAEAREKQAEELARRHEKREAELEAKOKVEEERRQTAOS
QOERRRRAAPQDYMAASODPRKRGHRSVOONNVEIAQOAEKARQOAEKRAEL
AKORAAEERQOALAAARRKAEAEKQAEELARRHEKREAELEAKOKVEEERRQTAEL
AAOROKALARQOAEKRAELAVKOKAEETERTABLAKORAAAEAKARQOAEERQTAEL

ARROAERAOALSLAKOKAETDREAESAKKRAKEEERROAOSOPORRRRAAODY
MAASORRRRRRRRSTLPAPSPSPDSSAYAPRALHNDWENDYEETPLDLEDEN
VSESVTSDKQPDNTELEHEKENDYEEIPLDLEDVSESVTSDKQPDNTELEHE
KRETVSLQPRAPRAQAATOLQAOAAQADAVSNTSALSDDAAQSQSLDITGAS
LIRHIAKSRADAEKNSVMSNTGYRBYASNOYRFSKSRTOIGIDRSISEMJOI
GGLVLYSDSOHTFDQSGKTFVQANLKGKRYANWYAVAGIGAGSLRSRTOQKA
NENRASICTGLIGMTLKNPEIYPSAGIRYSRLSADYKLGNDSVKSSMKTTLT
AGLDFAYRFRKGNLTVKPLLSAAYFANYGKGVNNGNSFVKKANNOQYSGAALLY
RVTLVNNGSTIKGKLEKOKSGQIKIOIR*

gene
complement(9703...9957)
/gene="NM00701"
complement(9703...9957)
/gene="NM00701"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAFA1118.1"
/db_xref="GI:7225929"
/translation="MTNPSVRRPAPRHSHSPSRHYHESGNLGLRVKKTLYIPSPPOKO
KTRINLKSRRHSRKSNGNPKTRNLTPEPNENRPLPFA"

BASE COUNT 2182 a 2461 c 2686 g 3156 t
ORIGIN

Query Match 4.4%; Score 192.4; DB 1; Length 10485;
Best Local Similarity 58.1%; Pred. No. 5,2e-34;
Matches 359; Conservative 0; Mismatches 256; Indels 3; Gaps 1;

QY 1199 aaggaagagtgatgacttaccagcaatcaacaaagcgagcggtgtgtat 1258
DB 8463 AAAACACACGCGACTTGGTATGTGATCAAAACATCAACAGCGCGGCTGT 8404
QY 1259 ttgaaggtatattttagctgtcgccttaaaaaaagaa---acgtgcaagcgcg 1315
DB 8403 TCAAAAGCGCATTTACACAGTCAAAAGTATTAATATGACATCCTTGGCT 8344
QY 1316 ttcatacagtgatgacgtaacgcttacttgaagaaagcggtgcaaacgacgc 1375
DB 8343 TTGATGTTGCGCGACGCAAAAAGTCTTGGCAAGTCAAAAATCCGAATG 8284
QY 1376 tgcctcaaatcgcaagcgacgctgtgttcaagcaaggggaaacaaagcg 1435
DB 8283 TGGCAAAATTCGCGCAAGGACATTTAGAAATTAACGCGCACAGCGCTTA 8224
QY 1436 tcaagcgtagcgagcgtaagtaacatccttagatcagcagcgagcgatca 1495
DB 8223 TAAAGTTCGCGCGAGGTATTCGATTAATCAACAGCGCGATGCGCAAAA 8164
QY 1496 aagcctttagaagcgtgtgttgcagcgagcgagcgagcgagcgagcgat 1555
DB 8163 AGGCTTTCCCGCAGTCGCGATTTGTCAAGTCAACAGCGCGATGCGCAAAA 8104
QY 1556 atcagttcaaccccgcaaaactctatcttgcgcttcgagcgagcggttga 1615
DB 8103 ATCAGATTAATCCCGATTAATTTCCGTTCCGTCGCGCTCTTTGATGCA 8044
QY 1616 ggcattcgcttcgctcccgccgattcaaatccagatgaagggcgatgata 1675
DB 8043 GCAATGACTTGAATTTGAACACATCCGCAAGCTGATGAAGCGCGCGCAT 7984
QY 1676 acgaatcaagaagaatccacgcttaccatcaagcaataaagatattacta 1735
DB 7983 ACAACACAGCGCGCTCCACATACGCTAAATCTTTGATTAACGATCCCA 7924
QY 1736 gcaatacaacaacttggatcaaaaagaatcttgcatacaagcttggtaga 1795
DB 7923 AAACATCTCTATCTATTAATCAATAAATATGATGATGATGATTAATTA 7864
QY 1796 aagatgcaacaaacga 1813
DB 7863 ACCGCGCTAGAAACCGA 7846

RESULT 21 AX044030 349980 bp DNA linear PAT 24-NOV-2000
AX044030/C Sequence 109 from Patent W00066791.
LOCUS AX044030
DEFINITION
ACCESSION AX044030
VERSION AX044030.1 GI:11342914
KEYWORDS
SOURCE
ORGANISM
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria meningitidis.
Neisseria meningitidis
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 349980
/organism="Neisseria meningitidis"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces.-seq 1: 1 to
349980 349980 bases-seq 108: 300001 to 649980 349980
bases-seq 109: 600001 to 949980 349980 bases-seq 110:
900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980
349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq
113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to
227325 17325 bases"
Location/Qualifiers
1. 349980

BASE COUNT 82523 a 82940 c 96712 g 87805 t
ORIGIN

Query Match 4.4%; Score 192.4; DB 6; Length 349980;
Best Local Similarity 58.1%; Pred. No. 9.7e-34;
Matches 359; Conservative 0; Mismatches 256; Indels 3; Gaps 1;

QY 1199 aaggaagagtgatgacttaccagcaatcaacaaagcgagcggtgtgtat 1258
DB 130037 AAAACACACGCGACTTGGTATGTGATCAAAACATCAACAGCGCGGCTGT 129978
QY 1259 ttgaaggtatattttagctgtcgccttaaaaaaagaa---acgtgcaagcgcg 1315
DB 129977 TCAAAAGCGCATTTACACAGTCAAAAGTATTAATATGACATCCTTGGCT 129918
QY 1316 ttcatacagtgatgacgtaacgcttacttgaagaaagcggtgcaaacgacgc 1375
DB 129917 TTGATGTTGCGCGACGCAAAAAGTCTTGGCAAGTCAAAAATCCGAATG 129858
QY 1376 tgcctcaaatcgcaagcgacgctgtgttcaagcaaggggaaacaaagcg 1435
DB 129857 TGGCAAAATTCGCGCAAGGACATTTAGAAATTAACGCGCACAGCGCTTA 129798
QY 1436 tcaagcgtagcgagcgtaagtaacatccttagatcagcagcgagcgatca 1495
DB 129797 TAAAGTTCGCGCGAGGTATTCGATTAATCAACAGCGCGATGCGCAAAA 129738
QY 1496 aagcctttagaagcgtgtgttgcagcgagcgagcgagcgagcgagcgat 1555
DB 129737 AGGCTTTCCCGCAGTCGCGATTTGTCAAGTCAACAGCGCGATGCGCAAAA 129678
QY 1556 atcagttcaaccccgcaaaactctatcttgcgcttcgagcgagcggttga 1615
DB 129677 ATCAGATTAATCCCGATTAATTTCCGTTCCGTCGCGCTCTTTGATGCA 129618
QY 1616 ggcattcgcttcgctcccgccgattcaaatccagatgaagggcgatgata 1675
DB 129617 GCAATGACTTGAATTTGAACACATCCGCAAGCTGATGAAGCGCGCAT 129558
QY 1676 acgaatcaagaagaatccacgcttaccatcaagcaataaagatattacta 1735
DB 129558 ACCGCGCTAGAAACCGA 7846

Db 1159 TTGATGTTACCGACGCAAAAAGTCGTTTGGCAAGTCGCAATCCGAATGCGCACAGAT 1218
 Qy 1376 tttccaaaatcggcaaaagcagcgtcgtgttccaaagcaaggaagcaagcgtcg 1435
 Db 1219 TCGCAAAAATCGGCAAAAGGACATTAAGAAATAAAGCGGACAGCGGTTAACCAAGGCGCAT 1278
 Qy 1436 taagctgggagcagcgtatcaatcctcctcagcagcagcagcagcagcagcagc 1455
 Db 1279 TAAAGTCGGCGAGTACGGTTATTCGAATCAACAGCGGATCCGCAAAAAGTCC 1338
 Qy 1496 aagcctttagtgaatcgcgtgtgtcagcagcagcagcagcagcagcagcagcagc 1555
 Db 1339 AGCGTTTCCCAAGTCGCGCATTTGCACGCGGACGCGGTACATTGGTATTAATGTTCAA 1398
 Qy 1556 atcagcttaccgcaaaactctattcgcgttgcgttgcgttgcgttgcgttgcgtt 1615
 Db 1399 ATCAGTTATATCCCAATACCTATATTCGTTCCGTGCGGCTGTTGATGCGCAATG 1458
 Qy 1616 ggcattcgttgcgttccacgcagcattcaaatcagcagcagcagcagcagcagc 1675
 Db 1459 GCAATGACTTGGACTTTGGAACACATCCGCAACGTGATGAAGCGCGCATTTGTCACAC 1518
 Qy 1676 acaatcaagaagaagaatcaccgttaccatcagcagcagcagcagcagcagcagc 1729
 Db 1519 ACAACAGACGCGCTCCACATCACGCTAACGGTAAATCTTAATTACCA 1572

RESULT 25
 AF235032 4659 bp DNA linear BCT 02-MAR-2001
 LOCUS Neisseria meningitidis Iga1 protease gene, complete cds.
 DEFINITION AF235032
 ACCESSION AF235032.1 GI:13183004
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Neisseria meningitidis.
 Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
 Neisseria.
 1 (bases 1 to 4659)
 Vitoovski, S. and Sayers, J.R.
 Degenerate specificity of *Neisseria meningitidis* Iga1 protease
 Unpublished
 2 (bases 1 to 4659)
 Vitoovski, S. and Sayers, J.R.
 Direct Submission
 Submitted (16-FEB-2000) Division MGM, University of Sheffield, F
 floor, Royal Hallamshire Hospital, Sheffield S10 2JF, UK
 Location/Qualifiers
 1. 4659
 /organism="Neisseria meningitidis"
 /strain="NMB"
 /db_xref="taxon:487"
 1. 4659
 /codon_start=1
 /transl_table=1
 /product="Iga1 protease"
 /protein_id="AAK15023.1"
 /db_xref="GI:13183005"

/translation="MKTKRKRIKINLSIFLAVATPYSEALVDRDVQIIPRAE
 AKKRFVATDLVYKNGKOGNIGALSIVPITDVSADVRKRLITLIDPOYAVSKV
 KDELSTYGHNGHLDVSDNENEYSVAQNDYENKWHHNGNRLITLIDPOYAVSKV
 TEVAPITPAGGVEYTKDNRESEVAVAGQFQFNYNNTLEDYMARVIACT
 PYOVNVTASNLNGLIGFQDNSKHSPEKLEVLSONALTNVILGDSSEPLAYDK
 QEKRVFLGAYDWAGYOKNSQWEMNITKKFPAKIKQPDNAGIKNGSHHNTYFG
 TNSKIGSTFVLAENEDKANNQONVFNEDNTLIDONINAGAGLFPKDYVYKIG
 NDITMGAGIDVTGKRYVQKRVNDRILAKIGKGLLEIENGVNQGOIKVGGTYI
 LNOQADAKKVOAQSQVIGSCTGLVNSQINPNKIXFGPGGRLDNGDLTPE
 HIRNVDGARVNNHTSASTITLTKSLTNPSLSYHSIQNDYDDDDSYTTPRR
 PIPOKDLITKNITRIALKSGSVNAPPEPENGVEENDWYFMGTQDEAKNNANHN
 NORTSGEFGEEGNGHNGALNNGKSAQRILTGTGTNGKISVQGVLLS
 GRPTPHARDPVNKSARDAHFSKNNEVFEEDDIINTFRAELIAYVQSAFSSGRNV
 SNITANITATDNAAKVNIGYKNGDEVCSVDYTVCTNGLSIDKALNSPGARDQINCN

BASE COUNT 1542 a 1074 c 1041 g 1002 t

Query Match 4.38; Score 190; DB 1; Length 4659;
 Best Local Similarity 61.0%; Pred. No. 1.6e-33;
 Matches 326; Conservative 0; Mismatches 205; Indels 3; Gaps 1;

Qy 1199 aaggaagaagtgatctgacattcaccagcaacatacacaagcagcagcagcagcagc 1258
 Db 1085 AAGACAAAGCGCACTTGTGTTGATCAACATCAACCAAGCGCGGCTCTGTTT 1144
 Qy 1259 tgaaggtaattacggttcgcctaaacaagaa---acgtggcaagcgcggcg 1315
 Db 1145 TCAAAAGCGATTTACACAGTCGCAAGGATTAATATACATCTGCGTACGCGGGA 1204
 Qy 1316 ttcatacagtgatggcagcagcagcagcagcagcagcagcagcagcagcagc 1375
 Db 1205 TTGATGTTACCGACGCGCAAAAAGTGTGTTGCAATCAAAAATCGCAAGCGCGCACAGAT 1264
 Qy 1376 tttccaaaatcggcaaaagcagcgtcgtgttccaaagcaaggaagcaagcgtcg 1435
 Db 1265 TGGCAAAAATCGGCAAAAGGACATTTAGAAATTAACCGCACAGCGCTTAACCAAGGCAAT 1324
 Qy 1436 tcaagctggagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1495
 Db 1325 TAAAGTCGGCGAGGATACGTTATTCGTAATCAACACCGCATGCGCACAAAAGTCC 1384
 Qy 1496 aagcctttagtgaatcgcgtgtgtcagcagcagcagcagcagcagcagcagcagc 1555
 Db 1385 AGCGTTTCTCCCAAGTCGCGATTTGCAAGCGGCGGATGCTATTAATAGTTCAA 1444
 Qy 1556 atcagcttaccgcaaaactctattcgcgttgcgttgcgttgcgttgcgttgcgtt 1615
 Db 1445 ATCAGTTATATCCCAATACCTATATTCGTTCCGTGCGGCTGTTGATGCGCAATG 1504
 Qy 1616 ggcattcgttgcgttccacgcagcattcaaatcagcagcagcagcagcagcagc 1675
 Db 1505 GCAATGACTTGGACTTTGGAACACATCCGCAACGTGATGAAGCGCGCATTTGTCACAC 1564
 Qy 1676 acaatcaagaagaagaatcaccgttaccatcagcagcagcagcagcagcagcagc 1729
 Db 1565 ACAACAGACGCGCTCCACATCACGCTAACGGTAAATCTTAATTACCA 1618

RESULT 26
 NMIG4771 1489 bp DNA linear BCT 14-MAR-1995
 LOCUS N.meningitidis iga gene (EN3771).
 DEFINITION X82468
 ACCESSION X82468.1 GI:732865
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Neisseria meningitidis.
 Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
 Neisseria.
 1 (bases 1 to 1489)
 Lomholt, H., Poulsen, K. and Kilian, M.
 Comparative characterization of the iga gene encoding Iga1 protease

TITLE	Gene structure and extracellular secretion of <i>Neisseria gonorrhoeae</i>
JOURNAL	Iga protease
MEDLINE	Nature 325 (6103), 458-462 (1987)
FEATURES	87115823
SOURCE	location/Qualifiers 1..4899

KAETHYFYGOYNGHNDVADKENEYRVENONNYPBKAKGASNLGRLEDYNNARBNKE
VTEVAPLAPTPACGGGLDTPYKDNKRNFSVPYRGAGKOLYBEKGYHOENEGYDRLD
SOATRYATAGTPIKIDINDOTMNTGELGFGHNRKQVSAEELKQASODALNTGVLG
DSGSPLEAFDOKQKQWVFLGYDYWAGGAKSMDWMTYKKEFAKIQHNDAGVYK
NGEHNKWTGTGNSHIGTAVALRYANNEGANNONGANTFEDNGLVYKQINOGAGLFF
KQDYVKAANDITWLGAGIDVADKQVWVQKYNNGRLAKIGKGTLEIKGTGNTGNOG
OLKYGDTVLINOKADADKQVAFQVGVLSRGVLYSLNOSHSMWLTGYSAGHY
DANGNDLFEHRIHNDGARIYNNHTDASITTLTGKSLITNPNSLASHIONDDED
DXYTYRRPRPIPOCKDLYKNNRYRYALKSGRLNAPPPENVAANNMIFNGYTOE
ARKNMMNNKRRIGDEGFEDEMGHGNALNENKSKQNFLLTGANLNGKI
SVTOGNVLISGRPPHARDPVNKSARKDAHSKNEVEFEDDMINRFKAELAVNO
SASFSGRNVSDITANITATDNARKVNLGKMGDEVCSVDYTGTCGTGNLSDALN
SPDARINGNANLONALVLAALMGKIOGQNSRVSLSNOSHSMWLTGYSAGHY
LADSHIHUNNASDAQSANKYHTIKINLSGNGHFHYLDLAKNLCDKLYVKSAGHY
OLHYONKTEPNOEGIDLEFASVODRSRLFVSLANHYLDLALRYITKTENGITRLY
NPYANGRPVKPAPSPANTASOAKOTOTDGAOLAKPONIVVAPSPQANAEALR
QOAKAEYKROOAAEAERYAROKDEAKRAELIKQOEPEAKRAELAKOAEERK
ARELAKOAEASHOANAKPKRRRRRALLPRPAPVFLDDYDADNBSSESGNLARY
IPRMGRELINDYEIPLLEDEBAEERROATOFHSKSRNRALISSEPSDASBY
STSDKHPODNTLEHKEVETAGIOPRAOQPTAOAODAVSTNTSALSADMASTOSI
LIDTGAYLTRHIAOKSRADAENSVMSNTGGRDYASAQYRFSKRTOTQIGIDRS
LSENNQIGVLTYSOHTFPOAGCKNTFVONLQSGYLYNDAMYVAGDIGDSLSR
LOTQOKANENRSTIGTLGNTLKI NOFEIYPSAGIRYSRLSADYKLGDDSVYSS
MAVTLTAGIDFAFRKYNLTVPKLSAAYFANYGKGVNNGKSPAYKADNOOYS
AGVALLIYRVNTLVNNGSITKGLKOLEKSGOLKIDIRF"

BASE COUNT 1617 a 1105 c 1079 g 1098 t
ORIGIN

Query Match 4.3%; Score 188.8; DB 6; Length 4899;
Best Local Similarity 61.0%; Pred. No. 3.2e-33;
Matches 345; Conservative 0; Mismatches 212; Indels 9; Gaps 2;

QY 1167 gaataatgagaataatcttccatttgaacaaaggaagtgataactacacg 1226
DB 1192 GAACAACGGCAAAAGTTACTTTCMAAGCAACGCA-----CTTGTGATTGATCA 1245
QY 1227 caacatcaacaaagcgcggttgaatttgaagtaatttaagcttcgcctaa 1286
DB 1246 AATCATCAACCAAGCGCGCGCGCGCTGTTTCAAGCGCATACACAGTCAAGCGCC 1305
QY 1287 aacaacagaa---acgtggcaagcgcggtcatalcatagtgagcagttac 1343
DB 1306 AATATATGACATCACTTGGTTAGTGCGGGGATGTGATGTTGCCGCAAAAGTCTGT 1365
QY 1344 ttggaaagtaaacggtcggtgcaaacgacgcctgtccaaatcgcaagcgacgtgct 1403
DB 1366 TTGGCAAGTCMAAAATCCCAATGCGCACAGATTGGCAAAATGCGCAAGGCACTTTGGA 1425
QY 1404 ggttcaagccaaaggaacaaaggtcgtcagcgtggcgacggttaagtcatt 1463
DB 1426 AATTAACGGCACAGCGCTTAACCAAGGCAATTAAGGTCGGGACGATACGGTTATTC 1485
QY 1464 agatcagcgcgagcagatcaagcgcaaaacagccttagtgaatcggttgcag 1523
DB 1486 GAATCAAAAAGCGCATGCGCACAAAAGTTTCAGCTTTCCTCCCAAGTGGCATTTGTCA 1545
QY 1524 cggcaggggagcggtgcaatgaatgcgataatcaatcaaccccgcaactatt 1583
DB 1546 CGACACCGGTACATGTTGTTAAATGTTCAATAGATTAACTCCGATTAATATTT 1605
QY 1584 cggcttcgcgagcgacgttgaattgaacgacgttcgcttcgccaacgcatca 1643
DB 1606 CGGTTTCGCTGGCGGTCGTTTGGATTCGATGCGCATGACATTGACTTTTGAACACATCC 1665
QY 1644 aaataccgatgaaggcgatgatgtaacacaaatcaagacaagaatccacgctac 1703
DB 1666 CAACGTGATGAGGCGCGCATGTTGTCAACCAACACAGACACGCTCCACAAATCAC 1725
QY 1704 catcacgacaataaagattacta 1729
DB 1726 ACTAAGGGTAATCTTTAATTAACA 1751

RESULT 31
LOCUS A12416 4899 bp DNA linear PAT 05-OCT-1994
DEFINITION Iga-protease gene precursor.
ACCESSION A12416
VERSION A12416.1 GI:640585
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 4899)
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS

PROCESS FOR THE GENETIC PRODUCTION OF PROTEINS BY USING OF
GRAM-NEGATIVE HOST CELLS
Patent: WO 8800238-A 12 14-JAN-1988;
Location/Qualifiers
1..4899
/organism="Neisseria gonorrhoeae"
/db_xref="taxon:485"
/gene="Iga protease precursor"
104..4702
104..4702
/gene="Iga protease precursor"
/codon_start=1
/transl_table=1
/protein_id="CAA01030.1"
/db_xref="GI:640586"
/translation="MKAKRFINAIISLIFLAVALTPESEALVRDQVYJFRPFAE
MKGFPEVATGDSVKNRSGONATGNSLNPMDPSVADNKRILATVDPQVAVSKHA
KAEVHFYGOYNGHNDVADKENEYRVENONNYPBKAKGASNLGRLEDYNNARBNKE
VTEVAPLAPTPACGGGLDTPYKDNKRNFSVPYRGAGKOLYBEKGYHOENEGYDRLD
SOATRYATAGTPIKIDINDOTMNTGELGFGHNRKQVSAEELKQASODALNTGVLG
DSGSPLEAFDOKQKQWVFLGYDYWAGGAKSMDWMTYKKEFAKIQHNDAGVYK
NGEHNKWTGTGNSHIGTAVALRYANNEGANNONGANTFEDNGLVYKQINOGAGLFF
KQDYVKAANDITWLGAGIDVADKQVWVQKYNNGRLAKIGKGTLEIKGTGNTGNOG
OLKYGDTVLINOKADADKQVAFQVGVLSRGVLYSLNOSHSMWLTGYSAGHY
DANGNDLFEHRIHNDGARIYNNHTDASITTLTGKSLITNPNSLASHIONDDED
DXYTYRRPRPIPOCKDLYKNNRYRYALKSGRLNAPPPENVAANNMIFNGYTOE
ARKNMMNNKRRIGDEGFEDEMGHGNALNENKSKQNFLLTGANLNGKI
SVTOGNVLISGRPPHARDPVNKSARKDAHSKNEVEFEDDMINRFKAELAVNO
SASFSGRNVSDITANITATDNARKVNLGKMGDEVCSVDYTGTCGTGNLSDALN
SPDARINGNANLONALVLAALMGKIOGQNSRVSLSNOSHSMWLTGYSAGHY
LADSHIHUNNASDAQSANKYHTIKINLSGNGHFHYLDLAKNLCDKLYVKSAGHY
OLHYONKTEPNOEGIDLEFASVODRSRLFVSLANHYLDLALRYITKTENGITRLY
NPYANGRPVKPAPSPANTASOAKOTOTDGAOLAKPONIVVAPSPQANAEALR
QOAKAEYKROOAAEAERYAROKDEAKRAELIKQOEPEAKRAELAKOAEERK
ARELAKOAEASHOANAKPKRRRRRALLPRPAPVFLDDYDADNBSSESGNLARY
IPRMGRELINDYEIPLLEDEBAEERROATOFHSKSRNRALISSEPSDASBY
STSDKHPODNTLEHKEVETAGIOPRAOQPTAOAODAVSTNTSALSADMASTOSI
LIDTGAYLTRHIAOKSRADAENSVMSNTGGRDYASAQYRFSKRTOTQIGIDRS
LSENNQIGVLTYSOHTFPOAGCKNTFVONLQSGYLYNDAMYVAGDIGDSLSR
LOTQOKANENRSTIGTLGNTLKI NOFEIYPSAGIRYSRLSADYKLGDDSVYSS
MAVTLTAGIDFAFRKYNLTVPKLSAAYFANYGKGVNNGKSPAYKADNOOYS
AGVALLIYRVNTLVNNGSITKGLKOLEKSGOLKIDIRF"

BASE COUNT 1617 a 1105 c 1079 g 1098 t
ORIGIN

Query Match 4.3%; Score 188.8; DB 6; Length 4899;
Best Local Similarity 61.0%; Pred. No. 3.2e-33;
Matches 345; Conservative 0; Mismatches 212; Indels 9; Gaps 2;

QY 1167 gaataatgagaataatcttccatttgaacaaaggaagtgataactacacg 1226
DB 1192 GAACAACGGCAAAAGTTACTTTCMAAGCAACGCA-----CTTGTGATTGATCA 1245
QY 1227 caacatcaacaaagcgcggttgaatttgaagtaatttaagcttcgcctaa 1286
DB 1246 AATCATCAACCAAGCGCGCGCGCGCTGTTTCAAGCGCATACACAGTCAAGCGCC 1305
QY 1287 aacaacagaa---acgtggcaagcgcggtcatalcatagtgagcagttac 1343

Db	1306	AAATAAATGACATCACTACTTGTTAGTGTCGGGGATTGATCTTGCCGACGCAAAAGATCGT	1365
OY	1344	tttgaatgataaaccggtctgtggcaaaagaccgcgtctccaaatctggcaaaaggaacgctctc	1403
Db	1366	TTGGCAATCTCAAAAAATCCGAAATGGGGACAGATTTGGCAAAAAATCGGCAAAAGGCATCTTGA	1425
OY	1404	ggttcaagcccaaaagggaanaaccgaagctcgttcagcgctgagcgacgtaaaatcatctt	1463
Db	1426	AATAAACGGCAGCGCGTTAAACCAAGGGGCATTTAAAGTCGGCGACGGTAGCGTTATCT	1485
OY	1464	agatcagcagcggtggcgatcacaagggcaaaaaacaacccitttgatgaatccgctgtgtcaag	1523
Db	1486	GAAATCAAAAAGCCGATGCGCACAAAAGTTTAGGCTTTCTCCCAAGTCGGATTGTGAG	1545
OY	1524	cggcagggggagcaggtgcacactgaaatgacccgaataatgafttcaaccgccgacaactctatt	1583
Db	1546	CGGACGGCGTACATTGGTATTTAAATAGTTCAATACAGATTATATCCGATTAACCTATATT	1605
OY	1584	cggcttgcggtggcgagcttggatcttgaaaggggcatctgccttgccttccacccgatatca	1643
Db	1606	CGGTTTCGTGGCGGTCTCGTTTGATGATGCCAATGCAATGACCTTGACTTTTGAACACATCCG	1665
OY	1644	aaatccgatcgaaggagcgatgatgtctcaaccacacatcaagacaagaatccaccgtttac	1703
Db	1666	CAACGTGATGTAAGGCGCGCGCATGTTGCACACACAAGACAGCCAGCGCTCCACAAATCAC	1725
OY	1704	cattacgagcaataagaatatctacca	1729
Db	1726	ACTAAAGGGTAAATCTTTAATTACCA	1751

RESULT	32
LOCUS	AR079202
DEFINITION	Sequence 5 from patent US 5965424.
ACCESSION	AR079202
VERSION	AR079202.1 GI:10005948
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified. 1 (bases 1 to 4899) <i>Ambrosius,D., Dony,C. and Rudolph,R.</i> Methods for making neisseria or hemophilus Iga protease and DNA encoding the proteases Patent: US 5965424-A 5 12-OCT-1999;
JOURNAL FEATURES	location/Qualifiers 1..4899 /organism="unknown"
BASE COUNT	1617 a 1104 c 1080 g 1098 t
ORIGIN	

Query Match	Similarity	Score	DB	Length
Best Local	Similarity 61.0%	Pred. NO. 3.2e-33		
Matches	345; Conservative	0; Mismatches	212; Indels	9; Gaps
QY	1167	gaatataatgagaataatttccttatttgcagaagaagaagtgaattgatactaccag	1226	
DB	1192	GAACAACGGACAAAATGTTACTCTTTGAAACACACGGCA-----CTTGGTATTGAATCA	1245	
QY	1227	caacatacaaccaaggcgcgcgcgttgtatttgaaggttaatttcaagtcgtcctctaa	1286	
DB	1246	AAACATCAACCAAGGCGGGCGGCGCTGTTTTCAAAAGCGATTCACACGTAATAAAGCGCG	1305	
QY	1287	aaacaacgaa---acgttgcagaagcgcgcgcgttcataatcagttgacgtaccgttac	1343	
DB	1306	AAATATATGACATCATCTTGTTGTTAGTGTCGGGGATTAATGTGTGCCGCGGCAAAAAAGTCT	1365	
QY	1344	tttgaagtaaacgcgcgttgcacaacgaccgcctgttccaaatctgcgcaaaagcagctgct	1403	
DB	1366	TTGGGCAAGACAAAATTCGATATGCGGACAGATTTGGCAAAAATTCGGCAAAAGCACATTGGA	1425	

OY	1404	ggttcaagcccaaaagggaacacccaagcttcgycacgcygggcgcgylaaagtcattc	1465
Db	1426	ATTATACGGCAGCAGCGCTTATACCAACAGGCATTTAAAGTCGGCGAAGGATATTC	1485
OY	1464	agatcagcaagcggaagatcaagcaaaaaacaagcccttagtgaatcgccttgcg	1522
Db	1486	GAATCAAAAACCCCATCCCGACAAAAGTTCCAGGCTTCTCCCAAGTCGGCATTTGAG	1545
OY	1524	cggcagagggaagcgtgtgcaactgtaatgcgalaatcagttcaaccccgacaactatct	1583
Db	1546	CGGACGGCGTACATTGGTATTAAATAGTTCAATCAGATTATCCCGATTAACCTTATTT	1605
OY	1584	cggcttcgcggcggaagcttgaattgaagggagcttcgcttgcgtccacggatca	1643
Db	1606	CGGTTTCCGGCGCGCTGTGTGGATGCGCAATGCAATGACTTTTGAACACATCCG	1665
OY	1644	aaataccgaatgaaggggcgatalgtlglcaaccccaatcaagaanaaatccacggttac	1703
Db	1666	CAACGTGATGAAGCGCCGCGCATTTGCAACCAACACAGACCAAGCCTCCACATAC	1725
OY	1704	cattacagcaataaagaattacta	1729
Db	1726	ACTAAGCGGTAAATCTTAAATTACA	1751

LOCUS	8393 bp	DNA	linear	BCT 18-NOV-1996
NGCOM1GN				
DEFINITION	N.gonorrhoeae gene for competence lipoprotein ComL.			
VERSION	248895			
KEYWORDS	ComL; competence lipoprotein.			
SOURCE	Neisseria gonorrhoeae.			
ORGANISM	Neisseria gonorrhoeae			
	Bacteria: Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
REFERENCE	1 (bases 1 to 8393)			
AUTHORS	Fusseneegger,M., Racius,D., Meier,J. and Meyer,T.F.			
TITLE	A novel peptidoglycan-linked lipoprotein (ComL) that functions in natural transformation competence of Neisseria gonorrhoeae			
JOURNAL	Mol. Microbiol. 19 (5), 1095-1105 (1996)			
MEDLINE	96249702			
REFERENCE	2 (bases 1 to 8393)			
AUTHORS	Fusseneegger,M.			
TITLE	Direct Submission			
JOURNAL	Submitted (20-JUN-1995) Martin Fusseneegger Dr., Infektionsbiologie, Max-Planck-Institut, fier Biologie, Spemannstr. 34, Tuebingen, BW, D-72076, Germany			
FEATURES	Location/Qualifiers			
source	1..8393			

CDS	<code>/db_xref="taxon:485"</code> <code>120..923</code> <code>/transl_table=1</code> <code>/product="Comf, competence lipoprotein"</code> <code>/protein_id="CA99076.1"</code> <code>/db_xref="gi:110783"</code> <code>/db_xref="SPTREMBL:Q50985"</code> <code>/translation="MKILLIVSLGLALSACATGTADKDAQITQDWSVEKLYLAEAQ</code> <code>ELNSNRYRAVKLYLEISRPFTSKHASODSDLDYAYAKDDEKALALAEIRRRRLK</code> <code>POHNMDYALIRGLVLEFNEDOSFLNTKIASODSDRDREKAREAYQAEALVEORPNN</code> <code>KYADATRAWVKLIDALGNGNMSVARYMKGAVIAANRAKKIIGSYONTRYVEESI</code> <code>ALAEAYKKLIDKLPQLADTRRVLETFNPFKSPELTHAQPDPMWRYWH"</code>			
terminator				
BASE COUNT	2550 a	2300 c	1886 g	1657 t
ORIGIN	<code>934..967</code>			
Query Match	4.3%	Score 188.8;	DB 1;	Length 8393;
Best Local Similarity	61.0%;	Pred.No. 3.5e-33;		

Matches 345; Conservative 0; Mismatches 212; Indels 9; Gaps 2;

Qy	1167	gaataatggaagaataatttccttatttgacaagaagaagtgatgacttaccag	1226
Db	4686	GAATCAACGACAAATATGTTACCTTGAAGACACAGGCA-----CTTGGTATTGAATCA	4739
Oy	1227	caatcaacacgaagcgagcggttctgtatttgaaggtaatttcaacgctcgctaa	1286
Db	4740	AAATCATACCAACGCGCGCGGCTTTTTCAGAGCATTAACAGTCAAGGCGC	4799
Oy	1287	aaacacga---acgtggcaagcgcggttcataatcagtgatgacgttac	1343
Db	4800	AAATATGACATCACTGTTGTTAGTGGGGATGATGTTCCGACGCAAAAAAGTGT	4859
Oy	1344	ttggaagaaagcggtgcaaacgacgccttcccaaatcgcaagcagcgtgt	1403
Db	4860	TTGGCAGTCAAAAATCCGATGGCGACAGATGGCAAAAATCGCAAAAGCACTTTGA	4919
Oy	1404	ggttcaagcgaagaaggaagaagcgtcgtaagcggtggcgagcgttaagtcatt	1463
Db	4920	AATTAACGGCACAGCGCTTAACCAAGGCAATTAAAGTGGCGCGCTACGGTTATTCT	4979
Oy	1464	agatcaagcagcgagcagcagcagcaaaaacgaagcctttagtgaaatcggtgtcag	1523
Db	4980	GAATCAAAAACCCGATGCCGACAAAAGTTCAGGCTTTCCCAAGTCGGCATTTGTGAG	5039
Oy	1524	cgagcaggggagcgttgcacgtgaatgcccataatcagttcaaccccgacaactattc	1583
Db	5040	CGGACGGGTACATTTGATTAATTAATTCATTAATTCATTAATTCATTAATTT	5099
Oy	1584	cgagcttcgcgcgagcgttgcatttgaagcggagcgttcgttcaccccgactca	1643
Db	5100	CGGTTTCGGTGGCGGTGTTTGGATGCCAATGGCAATGACTTTTGAACACATCG	5159
Oy	1644	aaatacgaatgaagggcgatgattgtcaaccacaatgaagaacgaatccacggttac	1703
Db	5160	CAACGTGATGAAGCGCGCATTTGACCAACACACAGACAGCGCTCCACATAC	5219
Oy	1704	cattacaggaataaagataactacta	1729
Db	5220	ACTAACGGTAAATCTTAAATTACCA	5245
RESULT	34		
LOCUS	AF012206	2992 bp	DNA linear BCT 30-JAN-2001
DEFINITION	Neisseria meningitidis strain 24024 Iga1 protease precursor (Iga)		
ACCESSION	AF012206		
VERSION	AF012206.1	GI:2529495	
KEYWORDS			
SOURCE			
ORGANISM	Neisseria meningitidis. Neisseria meningitidis Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.		
REFERENCE	Morelli,G., Malorny,B., Muller,K., Seiler,A., Wang,J.F., del Valle,J. and Achtman,M. 1 (bases 1 to 2992) Clonal descent and microevolution of Neisseria meningitidis during 30 years of epidemic spread Mol. Microbiol. 25 (6), 1047-1064 (1997)		
TITLE			
JOURNAL			
MEDLINE	98010345		
REFERENCE	Morelli,G., Malorny,B., Muller,K., Seiler,A., Wang,J.F., del Valle,J. and Achtman,M. 2 (bases 1 to 2992) Direct Submission Submitted (02-JUL-1997) Max-Planck Institut fuer Molekulare Genetik, Imnestrasse 73, Berlin 14195, Germany Location/Qualifiers 1. .2992 /organism="Neisseria meningitidis" /strain="24024" /db_xref="taxon:487"		
FEATURES			
SOURCE			

mat_peptide

/note="serogroup A, subgroup VI; isolated in East-Germany, 1985"

/gene="iga"

/product="Iga1 protease"

/gene="iga"

/allele="iga6"

/gene="iga"

/translatability=1

/product="Iga1 protease precursor"

/protein_id="AAC45789.2"

/translation="LAVALPYSEALVRDQVDYQIFRDEAFENKKEFFVATDLSVNKOGONINATSNVPIDESVADVRRLTIVDIPQAVSVKRVKDELSYHNHGHLDVSDENYRSVAQNDYERPNKMMHNGRLEDDYMMALNKFTEVAPAPAPASAGGYETKDRNRESEVRVAGTQREYNSRYNMTLSRYRYRIACTPIQDVAVYNSLNQBELIGFDNSKHHSPERLKEVLSQNALNTAVLDGSPPLADKQRMVFLSATYDMAGYQKNSQWEMNLYKREFADETRKQDNAGTTRKGGHMKTTGTSTAVRAGNERDANNQVTFEDNCTLVLDQINQAGAGLEPKGDTYKGAANNITWLDGADVADGKVVQVQKPNCDRLAKIGKGLTNGTGVNOGQILVKGADSNQKQVAFQVGVSGRGLVNSNOINPDLVYFGRRGLDANGNDLTFEHIRNVEGARIYHNHNTSFASTITLTKSLITAPQNLVYEIRNDYDDDDYGYRYRPIQDGLYKYNRYVALKSGSVNAPBNPQVYENMDVWEMGYTOEPKRNANNNKNNORISGSGFGEENKGNALINENKSNQNRFLITGCANLNGKISVTOGNVILSGRTPHARDVYKSSARKDAHFSKNNVEYFEDDMINRTFAETLVNOSASFSQNRVSDITANITADNKLVLGYKNGDEVCSVDYTYCTNDLSDKALNSFATQINGVNLSONALVLGAAALMGLOGQNSRVSLNQHSHKMLTGDSOVNHLSDSHIHLNNAQSDASANKYHTLRIHLSGNGHFHYLTHLAKNLGDKVLYVESAGHYQLHVDKTPGNEGLNLPASSVQDRSLVSLVNNHVDGALRYTITENGSLITRLVNPYAEINRRVYKPAESPATNVAQOKATOTDGAQIAKQNTIVVAPSPQAN"

BASE COUNT

983 a 684 c 639 g 686 t

ORIGIN

Query Match

Best Local Similarity 58.4%; Score 188.4; DB 1; Length 2992;

Matches 349; Conservative 0; Mismatches 246; Indels 3; Gaps 1;

Qy 1199 aaggaagaagtgtaattgtaacttaccagcaacataaacaagcgagcgttctgatt

Db 1036 AAGACACGCGACTTTGGTATTGATCAAAACATCAACCAAGCGCGGTCTGTTT

Oy 1259 ttgaaggttaatttaccggttcgtcctaaacaaga---acgtggcaagcgagcggtc

Db 1096 TCAAGGCGATTACACAGCAACAAAGTGCMAATTAATGCGCATCTTGCGTACGCGGCA

Oy 1316 ttcatatcagtgatgagcgttaccgttacttgaagaagtaacgagcgttggcaaacgcccgc

Db 1156 TTGATGTTGCCGACGCAAAAAGTCTTTGGCAAGTCAAAAATCCGAATGGCGACAT

Oy 1376 tgtccaataatcgcaaaagcagcgttgcattcaagcgaaggggaaacaaagcgtcgg

Db 1216 TCGCAAAAATCGCAAAAGCACATTGAATAATTAAGGCGACAGCGCTTAACCAAGGCAAT

Oy 1436 tcaagtgagcagcaggttaagcattcattagcagcagcagcagcagcagcagcagcagc

Db 1276 TAAAGTGGCGACGATGCTTATTTGTAATCAAAAACCCATTCAAAACCAAAAGTCC

Oy 1496 aagccttagtgaatcggttcgttcagcagcagcagcagcagcgttgcacgtgaatgcgata

Db 1336 AAGCTTTCCCAAGTGGCATTTGACGCGACGCGGTACATTTGATTAATTAATTAATTA

Oy 1556 atcagttcaacccgacaaactatttgggttcgcgagcagcagcagcagcagcagcagc

Db 1396 ATCAATTAATCCCGATTAATTTGTTCCGTCGCGCGGTGATGATGATGATGATGATG

Oy 1616 ggcattcgttcgttcacgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc

Db 1456 GCAATGACTTACTTTGAAACATCCGCAACGTGATGAAGGCGCGCATTTGCAACC

RESULT 38
NMIGAG80 1489 bp DNA linear BCT 14-MAR-1995
DEFINITION N.meningitidis Iga gene (NGC80).
ACCESSION X82479.1 GI:732883
VERSION Iga gene: Iga1 protease.
KEYWORDS Neisseria meningitidis.
SOURCE Neisseria meningitidis.
ORGANISM Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 1489)
AUTHORS Lomholt,H., Poulsen,K. and Kilian,M.
TITLE Comparative characterization of the Iga gene encoding Iga1 protease
in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
influenzae
JOURNAL Mol. Microbiol. 15 (3), 495-506 (1995)
MEDLINE 95302961
REFERENCE 2 (bases 1 to 1489)
AUTHORS Lomholt,H.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-1994) H. Lomholt, Dept. of Medical Microbiology,
The Bartholin Building, University of Aarhus, 8000 C Aarhus C,
DENMARK
FEATURES
source Location/Qualifiers
1..1489
/organism="Neisseria meningitidis"
/strain="NGC80"
/db_xref="taxon:487"
1..1489
/gene="iga"
1..1489
/gene="iga"
/codon_start=1
/transl_table=11
/product="Iga1 protease"
/protein_id="CA57862.1"
/db_xref="GI:732883"
/db_xref="SPTREMBL:O51172"
translation="ATDLSVKNKOGCIGNALSNLPMIPSVADVNRITLVTDPOA
VSKHVGDEISVYGHNGHLDVSNDEYRSVAQDYEPNKMHHNGRGLDYVMA
RINKFEVAPLAPTSAGSGVETKDNKSESVRAGAGTOEYNSRYMTELSRYR
VALAGTPQDVAVTNSLNOEGILGSDNSKPEIKREVLSONLTVAVIGDSGSP
LFAVDKOEKRVETAGYADYAGYQKSNQKMEKPEKPEDEKQKDNATKINGSEH
KKTGNSHIGSTAVRLANERNNGQNVTEENSTLVLDQINOGAGLDFKGDIT
VGINNDITWLAGIDVADCKVVMQVKNPNDRKATGKGLTNGTGVNOGLKVG
DETVLINQADADKRVQAFSVQIVSGRTLVLSNSQINPNLVFGFRGGRIDANGN
DLFEHINVDGCAIRLVNHTGHASITLTGKSLITPKTISHYI"
BASE COUNT 498 a 321 c 320 g 350 t
ORIGIN

Query Match 4.2%; Score 186.4; DB 1; Length 1489;
Best Local Similarity 60.7%; Pred. No. 9.4e-33;
Matches 323; Conservative 0; Mismatches 206; Indels 3; Gaps 1;

QY 1199 aaggaagaagtgatgatactaccgacacatcaacgaagcgcggttgattat 1258
11
Db 926 AAAACAACGCGACTTGGTATGTGATCAAAACAACGAAGCGCGGCTGTTT 985
QY 1259 ttgaaggttaatttaaggtctgccttaaaacagaa---aagtcgaagcgcgggcg 1315
11
Db 986 TCAAAAGCGGTTACACAGTAAAGTAAATGATCATCTTGGCTAGTGGCGGGA 1045
QY 1316 ttcaatcagtgatggagtcgacttacttgaaagtaaacggtggaacagcgcc 1375
11
Db 1046 TTGATGTGCGGACGCAAAAAGTGTGCAAGTCAAAATCCGAATGCGACAGAT 1105
QY 1376 tgcctcaaatcgcaagagcgctgctgttcaagccaaaggggaaacaaagctcgg 1435
11
Db 1106 TGGCAAAATCGCAAGGCACTTGAATAAAGCGACAGCGCTTAACCAAGGCAAT 1165

QY 1436 tcaagctggcgacggtlaaagtcattcttagatcagcagcgacataaagcaaaac 1495
11
Db 1166 TAAAGTCGCGACGCTAGCGATTCTGATCAACACACCGCGCAAAAAGTGC 1225
QY 1496 aagcccttgtaaatcgcgttgtaagcgagcgaggcggtgaacatgacgata 1555
11
Db 1226 AGGCTTTCGCCAAGTCGCGATGTTCAGCGGACCGCGTACATGTGATTAATFAGTTCA 1285
QY 1556 atcagttcaaccgcgaactatcttgcttgctggcgaggttgatgttgaaag 1615
11
Db 1286 ATCAGATTATCCCGAATACCTATATTTCGTTCCGTGCGGTGTGATGCCAATG 1345
QY 1616 ggcattcgttcgtccaccgcacatcaaatccagatgaagggcgagatgtcaac 1675
11
Db 1346 GCATGACTGATGATTGTAACACATCGGACAGTGATGAGCGCGCATGTGCAAC 1405
QY 1676 acatcaagaacaaagatccacgcttaccatcattacaggaataaagatattac 1727
11
Db 1406 ACAACAGCGCGACGCTCCACATCAGCTAACGSGTAAATCTTGTGATTAC 1457

RESULT 39
AF012203 2980 bp DNA linear BCT 30-JAN-2001
LOCUS Neisseria meningitidis strain Z2491 Iga1 protease precursor (iga)
DEFINITION gene, partial cds.
ACCESSION AF012203
VERSION AF012203.1 GI:2529489
KEYWORDS Neisseria meningitidis.
SOURCE Neisseria meningitidis
ORGANISM Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 2980)
AUTHORS Morelli,G., Molony,B., Muller,K., Seiler,A., Wang,J.F., del
Valle,J. and Achtman,M.
TITLE Clonal descent and microevolution of Neisseria meningitidis during
30 years of epidemic spread
JOURNAL Mol. Microbiol. 25 (6), 1047-1064 (1997)
MEDLINE 98010345
REFERENCE 2 (bases 1 to 2980)
AUTHORS Morelli,G., Molony,B., Mueller,K., Seiler,A., Wang,J.F., del
Valle,J. and Achtman,M.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-1997) Max-Planck Institut fuer Molekulare
Genetik, Innesstrasse 73, Berlin 14195, Germany
FEATURES
source Location/Qualifiers
1..2980
/organism="Neisseria meningitidis"
/strain="Z2491"
/db_xref="taxon:487"
/note="serogroup A, subgroup IV-1; isolated in Gambia,
1983"
1..2980
/gene="iga"
/product="Iga1 protease"
1..2980
/gene="iga"
1..2980
/allele="iga4"
1..2980
/gene="iga"
/codon_start=3
/transl_table=11
/product="Iga1 protease precursor"
/protein_id="AAC45786.1"
translation="LAVALTPESEALVRDDVDYQIFDFAEKNGKFPVATDLSVKN
KOGIGNALSNVPMIPDSVADVNRITLVTDPOAVSVKVKVGDSEISVGHNGHLD
VSNDEYRSVAQDYEPNKMHHNGRGLDYVMAARKFEVAPLAPTSAGSGVET
TKDKNRSSEVVRGACAGTOEYNSRYMTELSRATRYALAKPTPDQNVNLDQVGL
IGFGNSNHSPEKLEVLSONALNTAVLVDGSGPLPAPDKRNVFGLDADYVWAG
YKNSQWQMNVTYKKEFADEIKQRDNAGTIKYGEGHMTTGTNSHIGSTAVRLANGNER
GANNQNVTFENNGTLVLDQNIINOGAGLDFKGDYTYVGANNGITWLAGIDVADGK

mat-peptide

gene

cds

SOURCE Neisseria meningitidis Z2491.
ORGANISM Neisseria meningitidis Z2491
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.

REFERENCE 1 (bases 1 to 311321)
AUTHORS Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C., Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T., Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N., Holtroyd,S., Jagsels,K., Leather,S., Moule,S., Mungall,K., Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M., Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
TITLE Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
JOURNAL Nature 404 (6777), 502-506 (2000)
MEDLINE 20222556
REFERENCE 2 (bases 1 to 311321)
AUTHORS Parkhill,J.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-2000) Submitted on behalf of the Neisseria Sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT Notes:
Details of N. meningitidis sequencing at the Sanger Centre are available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).

FEATURES
source
1. 311321
/organism="Neisseria meningitidis Z2491"
/strain="Z2491"
/db_xref="taxon:122587"
/note="serogroup: A"
115. 120
129. 587
/gene="NMA0698"
129. 587
/gene="NMA0698"
/note="NMA0698, possible periplasmic protein, len: 152 aa; unknown, contains a probable N-terminal signal sequence and lies within a region of unusually low GC content"
/codon_start=1
/transl_table=11
/product="putative periplasmic protein"
/protein_id="CAB83985.1"
/db_xref="GI:7379425"
/db_xref="SPTREMBL:O9J0Q2"
/translation="MTPLKILFLFLFVFLAIIILFFPSSDIESFQNYQFEVYD KGPANVILVMKDGNGNFDKTIISGLVEYKEDNIIYSYIDGQFADSCYKPEI LYKTIINKNHNIIINSMEKNFLSEDKIMKTRMLADPKKCNIGTLD"
742. 1077
/gene="NMA0699"
742. 1077
/gene="NMA0699"
/note="NMA0699, pseudogene, probable ABC transporter protein, len: 336 bp; similar to C-terminal of many ABC transporters e.g. TR:O46973 (EMBL:U47048), mtfB, Escherichia coli microcin transport protein (707 aa), fasta scores: E(): 2.6e-19, 53.2% identity in 111 aa overlap. Contains PS00211 ABC transporters family signature. NMA0686 may be the remainder of this pseudogene"
/codon_start=1
/transl_table=11
/product="putative ABC transporter protein (pseudogene)"
868. 912
/gene="NMA0699"
/note="PS00211 ABC transporters family signature"
complement(1101..2327)
/gene="NMA0700"
complement(1101..2327)
/gene="NMA0700"
/note="NMA0700, possible ribonuclease BN, len: 408 aa; similar to SW:RBN_ECOLI (EMBL:LI9201), rbn, Escherichia

coli ribonuclease BN (EC 3.1.1.-), fasta scores: E(): 2.3e-28, 34.3% identity in 268 aa overlap. Longer than rbn at the C-terminus. Also similar to TR:085449 (EMBL:AF067083) Vitreoscilla sp. hypothetical protein (376 aa), fasta scores: E(): 0.45e-08 identity in 333 aa overlap. Contains a region similar to NMA0677, fasta scores: E(): 4.3e-08, 64.4% identity in 45 aa overlap. Contains hydrophobic, possible membrane-spanning regions, as does rbn"
/codon_start=1
/transl_table=11
/product="putative ribonuclease BN"
/protein_id="CAB83987.1"
/db_xref="GI:7379426"
/db_xref="SPTREMBL:O9JVT9"
/translation="MTFLORLQGLADNKCICAFAMFVRRFDEEVPQAASMTFTTL ALPVLTVMVAASITFPVDFRMSDPSVFNQTIYQGDAMFDTINAREQANRLTA TGSVMIVYSLMILRTIDMTFRINRVNSQRRMMQFLYMLLPGPSLGVISFM VGSVDALASGAPQMSGNRTNRATITTEPILLKLYRVPVAPARAPGALATA FLELTRSLFTYMGNFDDGYSITGAFNAVPTFLMLNLMTLVLGAVITSSLTWQ GAFFRGFDSKRGFDVLKILLDLDAQEGKALPQERRRHINMGYDGLLEKLA RHGYIYSGRGQWVLTGADSIENLEFLFVYRPLPVERDHVNAQVADAVAMPCLQTLN MTLAEFDAQAKKQOQS"
1376. 1385
/note="Core DNA uptake sequence: gccgctctgaa"
/label=DUS
complement(2417..2426)
/note="Core DNA uptake sequence: gccgctctgaa"
/label=DUS
2444. 2453
/note="Core DNA uptake sequence: gccgctctgaa"
/label=DUS
2486. 2743
/gene="NMA0701"
2486. 2743
/gene="NMA0701"
/note="NMA0701, possible pseudogene, len: 258 bp; shows weak similarity to part of SW:MRBA_ECOLI (EMBL:M99166), wrbA, Escherichia coli TTP repressor binding protein (197 aa), fasta scores: E(): 0.064, 29.3% identity in 82 aa overlap. Also similar to part of TR:085450 (EMBL:AF067083) Vitreoscilla sp. TTP repressor binding protein (fragment) (124 aa), fasta scores: E(): 6.6e-14, 56.6% identity in 83 aa overlap"
/codon_start=1
/transl_table=11
/product="hypothetical protein NMA0701 (pseudogene)"
complement(2755..2764)
/note="Core DNA uptake sequence: gccgctctgaa"
/label=DUS
2778. 2787
/note="Core DNA uptake sequence: gccgctctgaa"
/label=DUS
2861. 2864
2871. 3530
/gene="NMA0702"
2871. 3530
/gene="NMA0702"
/note="NMA0702, len: 219 aa; similar to TR:Q52925 (EMBL:Z50189), exsB, Rhizobium melioli putative regulator of succinoglycan biosynthesis (not a transcriptional regulator) (234 aa), fasta scores: E(): 3.1e-15, 32.7% identity in 205 aa overlap. Also similar to many bacterial hypothetical proteins e.g. SW:YBAX_HAEIN (EMBL:U33798), H11191, Haemophilus influenzae hypothetical protein (196 aa), fasta scores: E(): 0.79e-28 identity in 173 aa overlap"
/codon_start=1
/transl_table=11
/product="hypothetical protein NMA0702"
/protein_id="CAB83989.1"
/db_xref="GI:7379427"

Query Match 4.2%; Score 185; DB 1; Length 1525;
Best Local Similarity 60.0%; Pred. No. 2e-32;
Matches 327; Conservative 0; Mismatches 215; Indels 3; Gaps 1;

```
OY 1199 aaggaagaagtgaattgatacttaaccagcaacatcaacaaagcgcggtgtgtatt 1258
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 962 AAGACAACGGCAGCTTGGTATTGATCAAAACATCAACCAAGGCGCGGTCTGTTT 1021
OY 1259 ttgaggtaatttaacgtctcgccttaaaacacgaa---acgtgcaagcgcggtgc 1315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1022 TCAAAAGCGGTTACACAGTAAAGGTCCAAATTAATGACATCTTGGCTAGGTGGGGA 1081
OY 1316 ttcatatcagatgagtgagtaacgttacttggaaagtaaacggtggcaaacgacgc 1375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1082 TTGATGTTGGCGAGCGCAAAAAGTCTTTGGCAAGTCAAAAATCCGAATGCTGACAGAT 1141
OY 1376 tgtccaaatcggcaaaagcgagctgctgtgttcaagccaaggggaaacaaagctcgg 1435
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1142 TGGCAAAAATCGGCAGGCAATTAAGAAATTAACGGCAGCGCTTAACCAAGGCGCAAT 1201
OY 1436 tcagcgtggcgagcgtgaagatcattatagatcagcagcgagcgtatcaagcaaaaac 1495
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1202 TAAAGTCGGCGAGCGACAGCACTTATTCTGATCAAAAAGCGATTCAAAACCAAAAGTCC 1261
OY 1496 aagcctttagtaaatcggtgtgtcagcgcgagggagcggtgcactgtaattgcgata 1555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1262 AGGCTTCTCCCAAGTGGCGGCTTGTACGCGGCGATTCATTTGATTAATTAATTCTCA 1321
OY 1556 atcagttcaaccccgacaaatctcatttcgcttcgcgcgagcgttggatttgaacg 1615
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1322 ATCGATTAATCCCGATTAACCTTAATTTCCGTTTCCGCGGTGCTTGGATGGCAAG 1381
OY 1616 ggcattcgcttcgttcacccgcatlcaaatccagatgaagggcgatgattgtcaac 1675
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1382 GCATGACTTACTTTTGATGACATCCGCAACGTGATGAAGCGCGCATTTGTCAAC 1441
OY 1676 acaatcaagaagaatccaccgttaccattacaggaataagattattatacaacg 1735
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1442 ACAACACAGGCGCGCTCCACAAATCAATTAACGGCAATTAATTAATTAATTAATTA 1501
OY 1736 gcaat 1740
    |||
Db 1502 AAAAT 1506
```

RESULT 44
NMIGAF48 1489 bp DNA linear BCT 14-MAR-1995

LOCUS N.meningitidis iga gene (HF48).
DEFINITION X82480
VERSION X82480.1 GI:732875
KEYWORDS iga gene; Iga1 protease.
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis; beta subdivision; Neisseriaceae;

REFERENCE 1 (bases 1 to 1489)
AUTHORS Lomholt,H., Poulsen,K. and Kilian,M.
TITLE Comparative characterization of the iga gene encoding Iga1 protease in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus influenzae
JOURNAL Mol. Microbiol. 15 (3), 495-506 (1995)

MEDLINE 95302961
REFERENCE 2 (bases 1 to 1489)

AUTHORS Lomholt,H.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-1994) H. Lomholt, Dept. of Medical Microbiology, The Bartholin Building, University of Aarhus, 8000 C Aarhus C, DENMARK

FEATURES
Source Location/Qualifiers
1..1489 /organism="Neisseria meningitidis"

gene /strain="HF48"
CDS /db_xref="taxon:487"
1..1489 /gene="iga"
<1..>1489 /gene="iga"
/codon_start=1
/transl_table=11
/product="Iga1 protease"
/protein_id="CA57863.1"
/db_xref="GI:732876"
/translation="ATDLSVKNKOGONIGNALSNPMDIFSVADVNRRLTVLDPOYA
VSVKHVKGDEISVYGHNGHLDVSDNENRVSVAONDYEPNKMMHNGOERLEDYNA
RLNKPTEVAPLAPPTAGGGEVYKDKNRESEVRVAGATOPFNRSRYMTLSRAYR
VLAGTPEVDVNTSLNLDNBLGIFRDNSSHHSPKLFADIEIKORDAGTICKYGEHH
LFAVDQERRWFLGAVDYWAGYQKNSWDMNTRYKREFADIEIKORDAGTICKYGEHH
WKTGTNSHIGSTAVRLAGNERNAGNONTFENNGTLVLDOINTOGAGLFEKGYT
VGINNDITWLGAGIDVADGKRVVQWKPNNGRLAKIGKGTLEINGTGNQGLKVG
DGVTLNOKADSMOKVOAFSOGVIGSRGTLVNSSNOINDMLYFGRGRLDANGN
DLFEHRRVDEGARIVNHNHTGHTSTLTGKSLITPNLSLVHSI"

BASE COUNT 496 a 319 c 325 g 349 t
ORIGIN

Query Match 4.2%; Score 183.6; DB 1; Length 1489;
Best Local Similarity 60.3%; Pred. No. 4.3e-32;
Matches 322; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

```
OY 1199 aaggaagaagtgaattgatacttaaccagcaacatcaacaaagcgcggtgtgtatt 1258
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 926 AAAACAACGGCAGCTTGGTATTGATCAAAACATCAACCAAGGCGCGGTCTGTTT 985
OY 1259 ttgaggtaattttagcgcttcgccttaaaacacgaa---acgtgcaagcgcggtgc 1315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 986 TCAAAAGCGATTAACAGTCAAGGTAAATTAATGACATCACTTGGCTAGGTGGGGA 1045
OY 1316 ttcatatcagtgatgagtaacgttacttggaaagtaaacggtggcaaacgacgc 1375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1046 TTGATGTTGGCGAGCGCAAAAAGTCTTTGGCAAGTCAAAAATCCGAATGCGACAGAT 1105
OY 1376 tgtccaaatcggcaaaagcgagcgtgctgtgttcaagccaaggggaaacaaagctcgg 1435
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1106 TGGCAAAAATCGGCAGGCAATTAAGAAATTAACGGCAGCGCTTAACCAAGGCGCAAT 1165
OY 1436 tcagcgtggcgagcgtgaagatcattatagatcagcagcgagcgtatcaagcaaaaac 1495
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1166 TAAAGTCGGCGAGCGATCGGTATTCTGATCAAAAAGCGATTCAAAACCAAAAGTCC 1225
OY 1496 aagcctttagtaaatcggtgtgtcagcgcgagggagcggtgcactgtaattgcgata 1555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1226 AGGCTTCTCCCAAGTGGCGGCTTGTACGCGGCGATTTGATTAATTAATTAATTCTCA 1285
OY 1556 atcagttcaaccccgacaaactctatttcgcttcgcgcgagcgttggatttgaacg 1615
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1286 ATCAATTAATCCCGATTAACCTTAATTTCCGTTTCCGCGGTGCTTGGATGGCAAG 1345
OY 1616 ggcattcgcttcgttcacccgcatlcaaatccagatgaagggcgatgattgtcaac 1675
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1346 GCATGACTTGACTTTTGAAACATCCGCAACGTGATGAAGCGCGCATTTGTCAAC 1405
OY 1676 acaatcaagaagaatccaccgttaccattacaggaataagattattataca 1729
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1406 ACAACACAGGCGCACACCTCCACAAATCAACGCGTAAATCTTAAATTAATTAAC 1459
```

RESULT 45

NMIGAF96 1489 bp DNA linear BCT 14-MAR-1995
LOCUS N.meningitidis iga gene (HF96).
DEFINITION X82475
VERSION X82475.1 GI:732879
KEYWORDS iga gene; Iga1 protease.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2002, 10:28:37 ; Search time 27321.5 Seconds
(without alignments)
3350.209 Million cell updates/sec

Title: US-09-303-518D-649

Perfect score: 4374

Sequence: 1 atgaataacacgacgaacg.....aattagctacgcctgtaaa 4374

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

GenBank: 1: gb_da: 2: gb_hgt: 3: gb_in: 4: gb_om: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_un: 14: gb_vl: 15: gb_da: 16: em_fun: 17: em_hum: 18: em_in: 19: em_mu: 20: em_om: 21: em_or: 22: em_ov: 23: em_pat: 24: em_ph: 25: em_pl: 26: em_ro: 27: em_sts: 28: em_un: 29: em_vl: 30: em_hgt_hum: 31: em_hgt_inv: 32: em_hgt_other: 33: em_hgtc_inv:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
------------	-------	--------------------	----	-------------

1	4374	100.0	4374	6	A96616	A96616 Sequence 64
2	4374	100.0	4374	6	AX044007	AX044007 Sequence
3	4374	100.0	10074	1	AE002547	AE002547 Neisseria
4	4374	100.0	349980	6	AX044034	AX044034 Sequence
5	4366	99.8	4374	1	NME242535	NME242535 Neisseria
6	4266	97.5	4385	1	NME296227	NME296227 Neisseria
7	3954.8	90.4	4407	6	A96620	A96620 Sequence 65
8	3866.8	81.5	4350	1	NME296276	NME296276 Neisseria
9	3566.8	80.2	349061	1	NME224291	NME224291 Neisseria
10	3508	61.6	4355	6	A96618	A96618 Sequence 64
11	2693.2	29.1	13818	1	U32710	U32710 Sequence
12	1274.4	28.8	4317	6	AR157320	AR157320 Sequence
13	1260.8	28.6	4319	6	NLIGA3327	NLIGA3327 Neisseria
14	1249.4	28.6	4319	6	NLIGA3327	NLIGA3327 Neisseria
15	318.8	7.3	578	1	NLIGA3327	NLIGA3327 Neisseria
16	313.6	7.2	584	1	NLIGA3327	NLIGA3327 Neisseria
17	280.6	6.4	572	1	NLIGA3327	NLIGA3327 Neisseria
18	280.2	6.4	572	1	NLIGA3327	NLIGA3327 Neisseria
19	202.6	4.6	2992	1	AF012207	AF012207 Neisseria
20	197.8	4.5	2992	1	AF012206	AF012206 Neisseria
21	196.2	4.5	10485	1	AE002424	AE002424 Neisseria
22	196.2	4.5	349980	6	AX044030	AX044030 Sequence
23	195.6	4.5	1489	1	NMIGA771	NMIGA771 Neisseria
24	195.6	4.5	1489	1	NMIGA771	NMIGA771 Neisseria
25	195.6	4.5	1525	1	NMIGA771	NMIGA771 Neisseria
26	195.6	4.5	1525	1	NMIGA771	NMIGA771 Neisseria
27	195.6	4.5	1525	1	NMIGA771	NMIGA771 Neisseria
28	195.6	4.5	1525	1	NMIGA771	NMIGA771 Neisseria
29	195.6	4.5	1525	1	NMIGA771	NMIGA771 Neisseria
30	195.6	4.5	1525	1	NMIGA771	NMIGA771 Neisseria
31	193	4.4	1492	1	NGC001202	NGC001202 Sequence
32	193	4.4	2980	1	AF012203	AF012203 Neisseria
33	193	4.4	2980	1	AF012204	AF012204 Neisseria
34	193	4.4	2980	1	AF012205	AF012205 Neisseria
35	193	4.4	2983	1	AF012209	AF012209 Neisseria
36	193	4.4	2985	1	AF012209	AF012209 Neisseria
37	193	4.4	4639	1	AF035032	AF035032 Neisseria
38	193	4.4	311321	1	NMA322491	NMA322491 Neisseria
39	192.4	4.4	1525	1	NMIGA771	NMIGA771 Neisseria
40	191.4	4.4	1489	1	NMIGA771	NMIGA771 Neisseria
41	191.4	4.4	1525	1	NMIGA771	NMIGA771 Neisseria
42	191.4	4.4	2995	1	AF012210	AF012210 Neisseria
43	189.8	4.3	1489	1	NMIGA341	NMIGA341 Neisseria
44	189.8	4.3	1489	1	NMIGA341	NMIGA341 Neisseria
45	188.2	4.3	1489	1	NMIGA341	NMIGA341 Neisseria
46	188.2	4.3	1489	1	NMIGA341	NMIGA341 Neisseria
47	188.2	4.3	2980	1	AF012211	AF012211 Neisseria
48	188.2	4.3	5261	1	NMIGA771	NMIGA771 Neisseria
49	186.6	4.3	1489	1	NMIGA771	NMIGA771 Neisseria
50	186.6	4.3	1522	1	NMIGA771	NMIGA771 Neisseria
51	186	4.3	1489	1	NMIGA771	NMIGA771 Neisseria
52	186	4.3	1489	1	NMIGA771	NMIGA771 Neisseria
53	186	4.3	1489	1	NMIGA771	NMIGA771 Neisseria
54	184.4	4.2	1489	1	NMIGA771	NMIGA771 Neisseria
55	183.4	4.2	1525	1	NMIGA771	NMIGA771 Neisseria
56	175.2	4.0	11288	1	HEA1GAI66	HEA1GAI66 Neisseria
57	174.4	4.0	172325	6	AE002549	AE002549 Neisseria
58	174.4	4.0	172325	6	AX044035	AX044035 Sequence
59	174.4	4.0	349980	6	AX044034	AX044034 Sequence
60	166.2	3.8	5536	1	HEA1GAI66	HEA1GAI66 Neisseria
61	166.2	3.8	5536	1	HEA1GAI66	HEA1GAI66 Neisseria
62	166.2	3.8	5536	1	HEA1GAI66	HEA1GAI66 Neisseria
63	161.8	3.7	5088	1	HEA1GAI66	HEA1GAI66 Neisseria
64	151.8	3.5	5719	1	HEA1GAI66	HEA1GAI66 Neisseria
65	149	3.4	5096	1	HICA	HICA Sequence
66	149	3.4	5096	1	HICA	HICA Sequence
67	141.6	3.2	213464	1	SPSPACN	SPSPACN Sequence
68	140	3.2	4484	1	EC027614	EC027614 Sequence
69	138.6	3.2	37770	1	AX276281	AX276281 Sequence
70	135	3.1	4113	6	AX276281	AX276281 Sequence
71	135	3.1	48254	6	AX276281	AX276281 Sequence
72	134	3.1	4394	1	SF035656	SF035656 Sequence
73	134	3.1	5639	1	AF097644	AF097644 Sequence

Db	181	AAAGCAATGTTGACGTGGGGCGCAAGGATATGACGTTTACACAAAAAAGGGAGTTT	240
QY	241	gtcggaatacaagaacaaagcccgatgatgttctcgtgagtgctgcgtaacgc	300
Db	241	GTGGCAATTCATACACAAAAGCCCGATGATTGATTTTCTGTGTCTCCGTAAAGGC	300
QY	301	gtcgcgagctgtgtggcgatcaatataatgttgaagctgtgaacgaataacgcgcgtataac	360
Db	301	GTGGCGCATTTGCTGGCGGATCATATTGTGTGTAGGTGACCATTAAGGGGGGCTATTAC	360
QY	361	aacgttgatcttgtgtcggaagagaagaatcccgatcaacatcgttttaacttaaaatt	420
Db	361	AAAGGTGATTTTGGCGCGAAGAGAAAGAAATCCGATCAACATCGTTTACTTATAAAATT	420
QY	421	gtgaacacggaataattaaagcagggactaaagcgcatccttatagcgcgatcat	480
Db	421	GTGAACAGGAATAATTATAAAGCGAGGACATAAAGGCGATCCTTATGGGGGAGATTATCAAT	480
QY	481	atgcgcgctttgcataaacttqtlccagatgcagaacccgtltgaaatgacccgltatag	540
Db	481	ATGCCCGCTTTTGCATAAAATTGTGTACGAGAGCGAAGCCGTGGAAATGACCGATATAATG	540
QY	541	gatbggcggaataatataatgcatacaaatataactacccgtacccgtltcgtbaltggygcagc	600
Db	541	GATGGCGGGAATAATATATGATGACCAAAATAAATTAATCCCTGACCGCTGTTCTGATTGGGCGAGCG	600

660

Db	661	GCAAGTGGCTATCTTGGCTGDTGGTGGGCAATACCTTTGCACAAATGGATCAGTGGT	720
Qy	721	ggacagctcaacttaggtgaagtgaaaaaattaaacatagggccatagttttttaccacg	780
Db	721	GGACACGTCAACTTGTAGTAGTGAATAAATTAAACATAGCCCATATGGTTTATACCAAC	780
Qy	781	ggaggtctacttggagcaaatggctcccaatggttatctatctatgattgcccataagc	840
Db	781	GGAGGCTCATTTGGCGGACAGTGGCTCCAAATGTATTATATGATGCCCAAAAGCAAAAG	840
Qy	841	tgttcaattaaatgggtgtatctcacaaggggaaccccatatagaaanaagcaatgcttc	900
Db	841	TGCTTAAATTAATGGGTATGTCAAACGGGCAACCCCTATATAGGAAAAGCAATGCTTC	900
Qy	901	caagctggttcctgaagaatctggtctctatgataactcttgcctgtagatacccatcagta	960
Db	901	CACCTGGTTCGTAAGAGTGTGTTCTATATGATGAATCTATGGCTGAGATACCCTATTCAGTA	960
Qy	961	ttctagaacacccgtccaaatcgggaatatctctttaaagccgataataatgacacaga	1020
Db	961	TTCTAGAACACCGTCAAAATAGGAAATACTCTTTAAAGACGATAATATATGACACAGA	1020

Db	1021	AAATATCATGCCAATGACATGACAAATCTCTGACCTAATAGTTTAAACAACGACACCGTT	1080
Qy	1081	caattgtttaaibgttctcttattccgagacagcaagaagaaactbgttcttcaibgttcgaagt	1140
Db	1081	CAATTGTTTAATGTTCTTTATCCGAGACGACGACGAAACCGTTTATCATGCTGCAGGCT	1140
Qy	1141	ggtgtccaaacagtgatcgcgccagacatcgataataibgagaaatatcttccattattgcgaa	1200
Db	1141	GGGTGTACACAGTTATTCGACCCAGATCGAATTAATGAGAAATATTTCTTTATTTGACGAA	1200
Qy	1201	ggaagaagcgaaatgtaattacttaccagcaacatcaatcaagaibgttcggagaatatacttc	1260
Db	1201	GGAAGAAGCGGAATTGATTCCTTACCACGACACATCATCAAGGCTGAGAGGATTTATATTTC	1260
Qy	1261	caagagagatttaccggtctcgctcgtaaaataacgaacatbgttcgaagcgcggtgcattac	1320
Db	1261	CAAGGAGAGATTTACCGGTCTCGCTCGTAATAACGAACATBGTTCGAAGCGCGGTGCATTAC	1320

Db 1261 CAAGCAGATTTTACGGTCTCGCTGAAATTAACGAACTTGGCAAGCGCGCGCTTTCAT 1320
Qy 1321 atcagatgaagacagtaacgcttacttgaagaaagtaacgagcgtgcaaacagcgcctgtcc 1380
Db 1321 ATCAGTGAAGACAGTACGCTTACTTGGAAAGTAAAGCGCGTGGCAAAAGCAGCCGCTGTCC 1380
Qy 1381 aaatcgcaagacagcgtcgaagttcaagccaaaggggaaaccaaagcttcgaacgc 1440
Db 1381 AAAATCGGCAAAAGGACGCTGACGCTTCAAGCCAAAGGGGAAACCAAGCGCTCGATCAGC 1440
Qy 1441 gtcgggcaagcgttacagtcatttctgatacagcagcagcgaataaagcaaaaacaaac 1500
Db 1441 GTGGGCGAGGAGTACGATATTTGGATCAGCAGCAGCAGATTAAGCAAAAACCAAGCC 1500
Qy 1501 tttagtgaatcagcgttgtaagcggcaggggtacagtgcaacgtgaatgccaatacag 1560
Db 1501 TTTAGTGAATTCGCTTGTGTGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
Qy 1561 tcaaaccccgcaaaaactctatttcgcttcgagcgtgagcgttctgataaagcagcgt 1620
Db 1561 TTCAACCCCGCAAAACTCTATTTGGCTTTCGCGGAGCGCTTGGATTTAAACGCGCAT 1620
Qy 1621 tcgcttcgttcacccgcttcaaaatacagatgaagggcgagatgatacgaacacac 1680
Db 1621 TCGCTTGTTCACCGATTCATTAACGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
Qy 1681 caagcaaaagaatccacgcttaccatlaacagcaataaagatattgtacacacgcaat 1740
Db 1681 CAAGCAAAAGATCCACCGCTTACGATTAACGATTAAGATTAAGATTAAGATTAAGATTAAG 1740
Qy 1741 aacaacagcttgatagcaaaaagaaatcgcatacagcttggtttgagcagaagaa 1800
Db 1741 AAACAACAGCTTGATGAGCAAAAAGAAATTCGCTTACCAACGCTTGGCTTGGCGAGAAAT 1800
Qy 1801 acgacccaaaagcaagcggcgtcgaacttgcttaccagcccgccgcaagaagcgcac 1860
Db 1801 ACGACCAAAAGCAAGCGGCGCTCAACCTTTTACGAGCCGCGGAGAGAGCGGCAAC 1860
Qy 1861 ctgctgcttcgagcagcaaatlaaacgcaacatcaacgcaacaaacagcgaactg 1920
Db 1861 CTGCTGCTTTCGCGGAGAAATTTAAAGCGCAACATCAGCAAAACCGCAAACTG 1920
Qy 1921 ttttcaacggagac 1980
Db 1921 TTTTTCACGGGAGCAAC 1980
Qy 1981 aaaggggcatctcctcgcggggaatcgtlgtggaacacagcgtgatacccgcatct 2040
Db 1981 AAAGGGGCAATCTCTCGCGGGAATCGTGGGCAACGATGATCAACCGCACATTT 2040
Qy 2041 aaaggggaaacttccaaattaaagcagcagcagcgtgtgttccgcaatgtgtccaa 2100
Db 2041 AAAGGGGAAACTTCCAAATTAAGGCGGACAGCGGTGTGTTCGCGCAATGTTCGCAAA 2100
Qy 2101 gtcgaagcagatgtgcatcttgacatcaacgcccacagcagtttltgtgtcgcacgcat 2160
Db 2101 GTGAAGGCGGATTTGCAATTTAGCAATACGCCCAAGCAGTTTGTGTGTCCACCGCAT 2160
Qy 2161 caaaac 2220
Db 2161 CAAAGCCACACAACTGTACAGCTTGGAGCTGAGCGGTGTACAAATTTGTTCAGAAAA 2220
Qy 2221 accatlaacgagataaagtgtatgtcttcaatgtactaaagacacgcaatcaagcgaatgc 2280
Db 2221 ACCATTTACCGAGATTAAGTGTATTTCTTCAATGACTAAGACGCAATCAAGCGCAATGC 2280
Qy 2281 gactctgcagatcaacgcatcttaaatcacaagcgttgcaacacacacacacacacac 2340
Db 2281 GATCTTGGCGATCAGGCTATTAATCTCACAGGCGTTGCCACACTCAACGCGCAATCTT 2340
Qy 2341 aatgcaaatgagatacagcttatacagtcagccaaagcccaaccccaaaagcgaactc 2400
Db 2341 ACTGCAAAATGGGATACAGCTTATACGTCAGCAACGCGCAACCCCAAAACCGCAACCTT 2400

Qy 2401 agctcgttgagcaatgcccagcaacatllaalcaagccacatlaaaagcgaacacacacac 2460
Db 2401 AGCCTCGTGGCAATGCCCAAGCAACATTTAATCAAGCCACATTAAGCGCAACACATCG 2460
Qy 2461 gcttcgagcaatgcttcaatlaactaagcagcagcagcagcagcagcagcagcagcagc 2520
Db 2461 GCTTCGGCAATGCTTCAATTAATTAAGCAACGCGGTGTACAAAGGAGGAGGAGGAGGAG 2520
Qy 2521 ctctcggcaacgctlaagcaacgtaagccatctcgcactcaacggttaatgtctcccta 2580
Db 2521 CTTTCGGCAACGCTTAAGCAACGTAAGCATTCGCACTCAACGGAATGTCTCCCTTA 2580
Qy 2581 ggcgaataagcagatcttcaatctgaagcagcagcagcagcagcagcagcagcagcagc 2640
Db 2581 GCGGTAAGGCAATTTCTCATTTTGAAGGAGCGCTTTACCGGACAAATATAGCGGCGCG 2640
Qy 2641 aagatacagcatlaactlaaagacagcgaatlgacgctgcgcgaagcaagaaatla 2700
Db 2641 AAGGATACGGCATTAACACTTAAAGACAGCAATGAGACGCTGCGTCAAGCGCAATTA 2700
Qy 2701 ggcgaattaaaccttgagcaacgcccacatlaactcaatctcgcactacgcagatgag 2760
Db 2701 GGCATTTAAACCTTGACACAGCCACCATTTACTCATTTCCGCTTACGCGCAGATGCG 2760
Qy 2761 gcagggcgcaaacccgagcagtgagcagagaagcgccgcgcgcgttcgcgcgttcgcgc 2820
Db 2761 GCAGGGCGCAAAACCGGCAAGTGCAGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2820
Qy 2821 cgtccctatctacgcttcaacccgcaactcgttagatccgcttcaacagcgtgagc 2880
Db 2821 CGTTCCTTATTCGCTTACACCGGCAACCTTGGAGAAATCCGTTTCAACACCGCTGAGC 2880
Qy 2881 gtaacggcaaatlgaaagcgcaggaacatccgcttctgtcggagcacttctgagcagc 2940
Db 2881 GTAAAGGCAATTAAGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2940
Qy 2941 cgcagcgaacaaatlgaaagctlgcgggaagctcgggaagcacttcaacacttgcgcgaac 3000
Db 2941 CGCAGCGCAAAATTAAGCACTGCGGAAAGTTCCGAAGCACTTACACTTGGCGCTGAC 3000
Qy 3001 aatacggcaacgaaactcgaagcctcgaacaaatlgaaagcagcagcagcagcagcagc 3060
Db 3001 AATACCGCAACGAAACCTGCAAGCTTCAACATTTGAGGAGGAGGAGGAGGAGGAGGAGGAG 3060
Qy 3061 aaacgcgtgcgcaaaacttaatttcaacctgcaaaagcaacagcgtcgtgcgcgcgcgc 3120
Db 3061 AAACCGCTGTCGAAAACTTAATTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 3120
Qy 3121 tggcgttaccacatcctcgcgaagcagcagcagcagcagcagcagcagcagcagcagc 3180
Db 3121 TGGCTTACCAATCTATCTGCAAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3180
Qy 3181 caagagcttccgcaaaactcgaagcagcagcagcagcagcagcagcagcagcagcagc 3240
Db 3181 CAAGAGCTTTCGCAAACTGCGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3240
Qy 3241 ggcgaagccttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3300
Db 3241 GCGCAAAAGCTTGAAGCGGCTGATTTGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3300
Qy 3301 gttcgcgaacccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3360
Db 3301 GTTTCGCAACCGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3360
Qy 3361 gagaataaaacgggtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3420
Db 3361 GAGAAAAAAGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3420
Qy 3421 accgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3480
Db 3421 ACCGCGCGGCTACCAACCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 3480

Qy	841	tgcgtatratatgaggatattgcaaacgggcaacccctctatagaaaaagaaatggtctc	900
Db	841	TGcGTATrATatGGgATtTGCaAAAGGcCAACCCCTATATgGAAaAAAGCAATGGcTTC	900
Qy	901	caagctggttcgtlaaagatfgattctatgatagaatacctcttgctgtagataaccattcagta	960
Db	901	CAAGCTGgTtTcGTlAAAGATtGgTtCTATgATgAAATCTTgCTGGAGATACCAATTCAGTA	960
Qy	961	ttctacgaacacgaatgcaaaatgsgaaatcctctcttlaacgaacgaataaataatgycacaga	1020
Db	961	TTcTAcGAaCaGATGCAAAATGGGAATACtCTTTTAAcGAAGATATATATGGCCAGTA	1020
Qy	1021	aaaatcgaatccaaaatatgaaacaaatctctgcctaatagataaanaacgcgaacctt	1080
Db	1021	AAaATcGAATCCCAaACtTGAAcCAATCTcTGCCtAAATAGATTAaAAACGAACGCCTT	1080
Qy	1081	caattgitttaatglttccttattatccgagacagaagaagaacctglttalcattgctgcagtt	1140
Db	1081	CAATtGtTTATgTtTCTTTATTCcGAGACAGCAAGAAACCGTtTATCAATGCTGCAGGT	1140
Qy	1141	gggtccaaacagttatcgaccacgaacgaaataatggaataaataatctctcttattgacga	1200
Db	1141	GGGTCAaAGTtATGACCCCAcAGCTGAATATATGGAGAAATATTTCTTTATTGACGA	1200
Qy	1201	ggaaaaggcgaattgatacttacttaccagcaaatcaatcaaatgctgcgtagagattatcttc	1260
Db	1201	GGAAaAGCCGAATtGATACTTACCAcCAACATCAATCAAGGTGcTGAGAGATATATATTC	1260
Qy	1261	caaggaagatttccggtctcgcctgtaaaataacggaacttgcgaagcgcgcggttcat	1320
Db	1261	CAAGGAATtTtTACGGTCTCGCTCGTAaAATTAACGAACCTTGGCAAGGGCGCGGCTTCAT	1320
Qy	1321	atcagtgaaagacagtaaccttactcttggaagataaagcgcgtagcgaacgaacgcgctgccc	1380
Db	1321	ATcAGTGAAGAcAGTACCGCTTACTTGTGAAGATTAACGGCGTGGCAAAcGACCGCTGCCC	1380
Qy	1381	aaaatcggcaaaagcgcgctgcacgcttcaagccaaagggaacccaagctctgatcagc	1440
Db	1381	AAaATCGCAAAAGCCAGCGTCACAGTTCAAACCAAGGGGAaAACCAAGCGCTCGATCAAGC	1440
Qy	1441	gtcgcgacgagtgtaacgttcatcttcttggatacagcagcagcagataaagcgcaaaacaagcc	1500
Db	1441	GTcGGCAcAGGTGACGTCAATTTGGATTCAGCAAGCAGACGATTAAGCAAAAAAACCAAGCC	1500
Qy	1501	tttggtaaatctggcttggcgaacgagcgaagggttagcgbgcacatgaatgcccgaataatcag	1560
Db	1501	TTTgTAAATCCGGcTtTGcCAcGGCAAGGGGTAGCGTGtCAACTGAATGCCGATATATAG	1560
Qy	1561	ttcaaccgccgaacaaactctatttcgagcttctcgcgcgagcagcttggatttaaacogggcat	1620
Db	1561	TTCAACCCCGCAAAACtCTATTTCtCGGCTTTCCGGGGGAGCGTtTGATTTAAACGGGGCAT	1620
Qy	1621	tcgcttctcgattccacgcttattcaaaataccgataggaaggcggaatggttcaacacaat	1680
Db	1621	TCGCTTCtTGtTCCACCGCTATTTCAAATPACCGATGAAGGGGGCATTTGTCAACACCAAT	1680
Qy	1681	caagaacaaagaatcccaacgcttaccattcaacgagcaataaagatatgtcttacaacogggcat	1740
Db	1681	CAAGACAAAGATCCACCGTtTACCATTTACAGCAATTAAGATATTTGTCTPACCAACCGGCAT	1740
Qy	1741	aacaaacgcttgcgtatgacaaanaaagaatctgcctcaacagcttggcttgcgagagaagaat	1800
Db	1741	AAcAAACAGCTTGGATGACAAaAAAGAAATTTGCCCTCAACAGGTGGTtTGGCGAGAAAGAT	1800
Qy	1801	acgacgcaaaacgaacggcgagctcacaactgttttaccagcccgcgcgagaagaacggcacc	1860
Db	1801	ACGACCAAAAGCAAGCGCGCGCTCAACCTTTTtTACACACCCCGCGCAAGACACCGCACCC	1860
Qy	1861	ctcgtcgtcttccggcggaacaaatttaaacgcgacaatacgcgaacaaacgcgaacgt	1920
Db	1861	CTcGTcGTCTTCcGGCGGAaAAATTTTAAGCGCAATATACGCAAAcCAAGCGCAAACTG	1920

OY	1921	tttccgagcgagacacacacgcgcgcctacacatcgtttaaagacacattgtgtcgaa	1980
Db	1921	TTTTTCACGGCGACACCMACCGCACGCTCTCATTTTAAAGACCAATGGTGGCA	1980
OY	1981	aaagagggcatctccctcgcgvgggaatcgtgtgysgaacaagcctgataccgcacattc	2040
Db	1981	AAAGAGGGCATTCCTCGCGGGGAAATCGTGTGGGACACAGATGGATTAACCGCAATTT	2040
OY	2041	aaagcggaaacatcccaaatltaaagtcggaacgagcggtgtgttcccgcaatggttcgcaaa	2100
Db	2041	AAACGGGAAACTTCCAAATTTAAAGCGGACAGCGGGTGTTCGCCGAATGTTGCCAAA	2100
OY	2101	gttaaaagcgatttggattttgagcaatbaagcccaagcagttttgtgtgtgcacgcat	2160
Db	2101	GTGAAAGCGCAATTTGGATTGTGACATACGCCCAAGCATTTTTGTGTGTGCAACGGAT	2160
OY	2161	caaaagcacacaalcgtgtacacglttcgscacgtgacggygtcgtgacaatltgttcgaaaaa	2220
Db	2161	CAAAAGCCACACATCTGTATACACCTGTGCGACTGGACGGGTGTACAAATGTGTGAAAAA	2220
OY	2221	accaattacgacgaataaagcgtatcgtctcatcgtacctaagaaccgacatacagcgcaatgtc	2280
Db	2221	ACCATTAACCGACGATTAAGTGATTTGGTTGCTCATTTAGTAAGCCGATCAGGGGCAATGTG	2280
OY	2281	gattcttcgagatacagctcatcttaaatctcacaaggtcttgcacacgttaagcgaaactc	2340
Db	2281	GATCTTCCGATTCACGGCTCATTTTAAATCTCACAGGGCTTGCCACACTCATACGGCAATCTT	2340
OY	2341	agtgcaaatltgacgatatacagcttatacagctbaagccaanaacgcaccaaccaaacgysaacctt	2400
Db	2341	AGTGCAATATGGCGATACACGTATTAACAGTCACACACACMACGCCACCCCAAAAGGGCAACCTT	2400
OY	2401	agctctgtggaacatgtgcacgaagaacatttaatacgaagccacaftbaaagcgaacaacatcgt	2460
Db	2401	AGCTCTGTGGGCAATGGCCCAAGCACATTTAATCAAGCCACATTAATTAAGGGCAACACATCG	2460
OY	2461	gtctcgagcgaattgtcttaattlaactlaagcgacacgcgcgtacaaaacggaagtcgtacg	2520
Db	2461	GCTTCGGGCAATGCTTATATTAACTTAAGGACACACGCGTACAAAGCGCAAGTGTGACG	2520
OY	2521	ctttccgcgcacaagctaaaggaanaacgftaagccattccgcacctaacaggtaatgtctcccta	2580
Db	2521	CTTTCGGGCAACGCTTAAGGCAAAACGTAAAGCCATTCGCACTAACAGTAATGTCTCCCTTA	2580
OY	2581	ggcgaataagcagatattccatttttgaagaacagcgctttacgcgcgaataacgaagcgcgcc	2640
Db	2581	GCCCATTAAGCAGATATTCATTTTGAAGACGACGCGCTTACGGGCAAAATCAGCGCGCGC	2640
OY	2641	aagataacgcatacacttaaaagaacgaaatgtagcgtgtgcgtcagcaagyaalta	2700
Db	2641	AAGATATACGGCATATTCATTAAAGACAGAGATGAGCTCTCGCTCAGGCAACGGAAATTA	2700
OY	2701	ggcaattttaaacccttgcacaacgcacacaatacactcaattccgcctatcgcacagatgtcg	2760
Db	2701	GGCAATTTAAACCTTTGACAAACGCCACCATTTACACTCAATTCGGCTTATGCCACAGATCG	2760
OY	2761	gtcagggcgccaaacacgycasgtgtgacagatgtgcgcgcgcgcggttcgcgcggttctgcgyc	2820
Db	2761	GCAGGGCGGCAAAACCGGCACTGTGCACAGATGCCGCCGCGCGCTTCGGCGGCTTCGGG	2820
OY	2821	cgttccctatahatcgttlaacacgcgaactctgttgaatctccggtttcaaacgctltgacg	2880
Db	2821	CGTTCCCTATTATTCGTTTAAACGCCCAACTGTGGGTGAATCCCGTTTAAACACCGCTGACG	2880
OY	2881	gttaaacgcaaatgtgaacggttcagagaacatctccgtttatgttgaactcttcgagctac	2940
Db	2881	GTAAACGGCAAAATTTGAACGGGTCAAGGGAACATTCGCTTATGTGGAACACTCTTGCGGTAC	2940
OY	2941	cgacagcgacaacattgaagctgtgcgysaaagttccgaagcacttacaacttggcgtgtcaac	3000
Db	2941	CGCAGCGCAAAATTTGAAGCTGTGGCGAAAGTTCCGAAGGACTTTACACTTGTGGCGGTACAC	3000
OY	3001	aataccgyscaacgaacctgtgaagccttcgcgaacaaltgaacggtgtgtgtgaaagaaaagcaac	3060

[illegible]

Db	4081	GCATTCAAACCGCTACCGCGGCGGAGATTAAAGCAGATTATTCATTCAAAACCGGCGCAAC	4140
Qy	4141	attccatacaagcctatttgaagcgtgtcctataacgaatgcggttggggcaagtccga	4200
Db	4141	ATTTCACATACGCGCTATTATTAGCCTGTGCCTATACCGATGCGCGCTTCGGGCAAAAGTCCGA	4200
Qy	4201	acaagcgtcataacacccgctgatttgcctcgaagatttcgcacaaacccgaagtcggaattg	4260
Db	4201	ACACGCGTCATACCGCGCTATTGGCTAGGATTTTCGGCAAAACCGGACGAGCGGAATAGC	4260
Qy	4261	ggcgtlaaaagccggaatcaaaagtttcaagctgtccctcaagtcgcgcgcgccaagaagc	4320
Db	4261	GGCGTAAACGCCGAATCAAAAGGTTTCAACGCTGTCCCTCCACGCTGCCGCCCAAAAGC	4320
Qy	4321	ccgcaactcgaagcgcgaacacacagcgcgcgcgcctcaaatagactacgcctgtgtaa	4374
Db	4321	CCGCACTCGGAAGCGCAACACAGCGCGGCGCATCAAAATTAGGCTACCGCTGGTAA	4374

RESULT	3	AE002547	10074 bp	DNA	linear	BC1 25-MAY-2000
LOCUS		AE002547				
DEFINITION		Neisseria meningitidis serogroup B strain				
ACCESSION		AE002547				
VERSION		AE002547.1				
KEYWORDS		GI:7227244				
SOURCE		Neisseria meningitidis				
ORGANISM		Neisseria meningitidis				
		MC58				
		MC58				
		section 189 of 206				

ORGANISM
Neisseria meningitidis muc8
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE
1 (bases 1 to 10074)
Tettelin, H., Saunders, N.J., Jeffries, A.C.,
AUTHORS

Complete genome sequence of *Neisseria meningitidis* serogroup B strain M58

JOURNAL MEDLINE PUBMED REFERENCE	SCIENCE 287 20175755 10710307 2 (bases 1 to 10074)	1809-1815 (2000)
---	--	------------------

AUTHORS
Tetzelin, H., Saunders, N.J., Helleberg, J., Jettles, A.C.,
Netelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Cleek, A., Parksey, D.S., Blair, E.,
Clifton, H., Clark, E.B., Cotton, M.D., Uiterback, T.R., Khouri, H.,
Qin, H., Vamthekyan, J., Gill, J., Scarlato, V., Massigalin, V.,
Pappa, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappelli, R. and Venter, J.C.

TITLE
Direct Submission

JOURNAL
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA

FEATURES	source	Location/Qualifiers
		1..10074
		/organism="Neisseria meningitidis MC58"
		/strain="MC58"
		/db_xref="taxon:122586"
		/note="serogroup: B"
gene		704..1375
		/gene="NM1983"
CDS		704..1375
		/gene="NM1983"
		/note="hypothetical protein; identified by Glimmer2; putative"
		/codon_start=1
		/transl_table=11


```
/product="hypothetical protein"
/protein_id="AA042311.1"
/db_xref="GI:7227245"
/translation="MPPDKYKLSLKENIFLAKKVAIVAOIHNLSPRENCOTLLQEQI
IHRKNASALSDIOTILNKRAYOYVISHIENGDEVDLSLKKINNIYAKDSDIAPG
DPTGSGVYGLDGSRAHPNPKVEIEMARVLONIGQSGSTTEAARMLKVMOROVF
MCKNRKATLFLFANGLMAGCGCILEISEMOPNFKLSAFRRSDDDIDISFYQNC
ISGID"
1537..2548
/gene="NMB1984"
/Note="This region contains an authentic frame shift and
is not the result of a sequencing artifact; similar to
sp:Q00840 GB:211857 PID:45066 percent identity: 97.40;
identified by sequence similarity; putative; IS1106
transposase, authentic frameshift"
complement(2864..7237)
/gene="NMB1985"
complement(2864..7237)
/gene="NMB1985"
/Note="similar to GP:1408569 percent identity: 76.08;
identified by sequence similarity; putative"
/codon_start=1
/translation_table=11
/product="adhesion and penetration protein"
/protein_id="AA042312.1"
/db_xref="GI:7227246"
/translation="MKTDRRTTETHRKAPKTRIRFSPAYATLCSFELPQANAH
TYGINVQYRDFAENKGFPAVGAKDIEYVNGKELVGSMTKAMIDFVSVNRGVA
ALVGDQYIVSAVANGYNNVDGAEGRNDQHRFYKIVKRNRYAGTGHYGCDDYH
MPRLHFTVDIAEPMNTSYMDGRKYIDQNNYPRVATIGARQYMRSDDEPNRRSSY
HISAYSLVAGTGNTPAONGSGGTVMIGSEKIKHSYGLPTGSGFSGSGSMFYDA
OKOKMLINGVLOTGNPYIGKSGNFOLRVDYDELPADTTSVEYEPPONKCYFND
DNNGTKINAKHEHNSLPNRKTRTYVNLVNSLEAPRPVYHAAGVNSYPRILNG
ENISFIDEKGEILTSNTNCGAGLYPGLDPTVSENNETMGTGKISSEPTVTK
VNGVANDRISKIGETLHYOAKGENGSIYSGDTYITLQDQDDKQKQASEILYS
GRTVOINADQNPDKLYFGFGRGLDNGSLSHRLOQNDDEGMILYQPAEHRITLLS
VTTGNKDIAVTGNNNSLDSKKEIAYNGMFGKDDTKTGRNLVYQPAEHRITLLS
GNTNLNGNTQNGKLFSGSRTPHAYVNLNDHWSQKEIPRGELTSDMTGLNFKVEK
TITDDVVIASLTFTQDISGNVDLADNAHLMTGLATLNGLSANGTRVYVSHNAQNG
NLSLVGNAQVETQDNLNGNTSAGNASFNLSGONSLSGNAKNAVSHSLNG
NLSLADKAVGFESSRFTGOISGGKDTALSKDESELTLSGELNMLNDNTYITLNS
AYRHDAAGQVTSATDAPRRSRSRSLSTLPTPSVSRNTLTVNKGKNGCTFR
FMSELFGYRDKLKLAESSEGTTLVANNNGNPASLEQIVTEKDKNPSEINMT
LQENHYDAGKMRVOLI RKDGERLHPVKEDELSBKLAEKAKKQAEKNDSDALI
AAGRDVETKESVAEPARQAGENGIMQAEERKRVQADKDTALQKQAEATRAITI
AEPRARVARDLPOLQPOPOPOQROLDISRYVNSGISEPSATLNSVFAVQDELREVA
EDRRNAVWTSIGIDTHHYSQDFRAYROOTDLROIGMNLISGVRGILFISNRENT
FDGIGNSARLARGAVGQYIDRFYIGISAGAGSSGSLGIGKIRRLARLHICIO
ARYRAGEFGGLIEPHIGATRYEVOKADYEVENVNTTGLAIPRYRAGIKADYSKPA
OHSITPYLSLSTYDASGKVRTRVNTAVLAQDFCKTBSAENGVNNEIKGFTLSHAA
AARGPOLEQHSGLKIGYRW"
complement(7371..7802)
/gene="NMB1986"
complement(7371..7802)
/Note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/translation_table=11
/product="hypothetical protein"
/protein_id="AA042313.1"
/db_xref="GI:7227247"
/translation="MTENMDKRAKRAVETVYSPELVBOITLSDVYQVIMACFISGP
LPPSDLAQINDIISGADRIMAKAEQVAREHTIROQDTNRKQSLTGLFSAILIL
LFAVFLWSGYPTAASLAGTVALAGAFVIGRSHDQKN"
8087..9433
/gene="NMB1987"
complement(8087..9433)
/gene="NMB1987"
/Note="similar to GB:1410328 SP:P25522 PID:290554 GB:U00096
PID:2367268 percent identity: 69.20; identified by
sequence similarity; putative"
```

```
/codon_start=1
/translation_table=11
/product="thiophene and furan oxidation protein ThdP"
/protein_id="AA042314.1"
/db_xref="GI:7227248"
/translation="MSDNPVTIIVAVTARGGVGIVIRISGNLLPMAOLGKTPKP
RTATYADFPTDGOAIDSGILFFPAAPSTGSDVLELOGHGPVYMALNRCLTELG
ARLAEPDEFTRKRAFLNDKLDLADAEVADLIIDSSSARLARLRSKDGFSRHGLV
EDLITLMLVEATLDFPEEDIDFLERADNRKIDGRKRVDDVLAHQGALLRGLN
VVLVGAENVGRSSLNLNLADEVALVITVDLGTTRDVRERLIDVYVNI VDTAGRE
TDDVERIGIERSKRAVSEADVALVVDREGLNERTALILALPEELKRIEHSKD
LHAHAAGFCTGAEVYALSAKTGDLIDLKTLILAEACMOGSEBCLFARIRHNAL
KAADESLALGNNHDIELFAHLRLAIVACGEITGEFTADLLDLVIFSRPICGK"
BASE COUNT 2307 a 2467 c 2721 g 2579 t
ORIGIN

Query Match 100.0%; Score 4374; DB 1; Length 10074;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgaaacacgcgacgaacgcgacgaacgcgacgaacgcgacgaacgcgacgcgc 60
DB 7237 ATGAAACACGCGAACGACGACGACGACGACGACGACGACGACGACGACGCGC 7178
QY 61 atcgcttcctgcctgcctgcctgcctgcctgcctgcctgcctgcctgcctgc 120
DB 7177 ATCCGCTTCCTGCCTGCCTGCCTGCCTGCCTGCCTGCCTGCCTGCCTGCCTGC 7118
QY 121 tggcgagacacacacacacacacacacacacacacacacacacacacacacac 180
DB 7117 TGCGCGGACACACTTATTTGCGCATCACTACCACTACCTACCTACCTACCTACCT 7058
QY 181 aaagcaagcttgagctgcgagcgaaagatactgaagcttacaacaaagcgagctg 240
DB 7057 AAGGCAAGTTCAGCTGCGGCGGAAAGATATTTAGGTTTACACAAAAGGCGACTTG 6998
QY 241 gtgcgcaaatcaatgacaaagcccgatgattcttcgtgctgctgctgctgctg 300
DB 6997 GTGCGCAATCAATGACAAAGCCCGATGATTTCTGCTGCTGCTGCTGCTGCTGCTG 6938
QY 301 gtgcgagcattgctgagcgaatatacttgagcgtgagcgaatatacttgagc 360
DB 6937 GTGCGGCGCATTCGTCGCGGATCAATATTTGTCGCGGCGGCGGCGGCGGCGG 6878
QY 361 aacgttgatttgctgctgagcgaagaaatcccgatcaacatcgcttactataaat 420
DB 6877 AACGTTGATTTTGGTGGCGGAAGAAATCCGATCAACATCGTTTACTTATAAAT 6818
QY 421 gtgaaacgaataataataaagcagcgacgaagcgaatccctctatggcgagcga 480
DB 6817 GTGAACGGAATTAATTAAGACAGGCGACTAAGCGCATCTTATGCGGCGGATTCAT 6758
QY 481 atgcgcgcttgcaataattgtcacagatgcgaacacgtgtgaagacgaattatg 540
DB 6757 ATGCGCGCTTGCAATAATTGTCAACATGCAACACTGTGTAAGTACACGATTATG 6698
QY 541 gatggcggaatatatgcatacaataataccctgacgcgtgctgctgctgctgctg 600
DB 6697 GATGGCGGAATATATGATCAATAATTAATTAATTAATTAATTAATTAATTAATTA 6638
QY 601 aggcgaatttgcaatcgcgaatgaagcgaagcgaatgaagcgaatgaatgaatga 660
DB 6637 AGGCAATATGCGCGATTCGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 6578
QY 661 gcaagtcgcatcttcgctgcgttgctgctgctgctgctgctgctgctgctgctg 720
DB 6577 GCAAGTCGCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6518
QY 721 ggcacagtcacactgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 780
DB 6517 GGCACAGTCAACTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6458
```


Dp	289011	AAAGCAAGTTTGGCAGTCGGCGGCAAAAGATTTAGGTTTACAAACAAAAGGAGGTTG	288992
Qy	241	gtcggaataatagacaanaagccccaatgaattttctgtgtcgtctaacgc	300
Dp	288931	GTCCGCAATTAATGACAAAAGCCCCCATGATTTGATTTTCTGTGGGTGCGCTTAACGGC	288932
Qy	301	gtgcgcgcatctgtgtggcgacaaataatctgtgacgtgtgcacataacgcgcgtataac	360
Dp	288931	GTGGCGGCAATTGGTGGCGCATATATTGTGTGGCTGTGCACATAACGGCGGCTTAAC	288932
Qy	361	aacgttgatttgtgtgggaagaaatccccaatacatcgttttacttaataatt	420
Dp	288931	AACGTTGATTTTGGTGGCGGAAGAAGAAATCCGATTAACATCGTTTACTTTTAAATT	288772
Qy	421	gtgaacgggaataattataaagcagggaactaaagccaaccttaatgtgcgcgcatcat	480
Dp	288771	GTGAACGGGAATTAATTTAAAGCGAGACATAAGGCCATCTTATGGCGCGATTTATCAT	288712
Qy	481	atgcgcgcttgcataaatttgcacagatgcgaacccgttgaatgacggttatatg	540
Dp	288711	ATGCGCGCTTTGTCATTAATTTTGTCAAGATGCAGAACCTGTGTAANTGACCACTTATATG	288652
Qy	541	gatggcggaataatatacgaatcaaaataatccctbaacggtttcgtatgtggcgagc	600
Dp	288651	GATGGCGGAATATATCGATCAAAATATATACCTTACCGGTGTCTATTTGGGCGAGCG	288592
Qy	601	agccaataattggcgactgaactgaagaatgagccaaccttaatgtgcgcgcatcat	660
Dp	288591	AGGCAATATTTGGCATGTGTGATGAAGATGACCCCAATTAACGCCAAGATTCTATTCATATT	288532
Qy	661	gcaagtcgtaattcttgcgtctgtgtgtgcgaataaccttgcacaaatgcatcaagtcgt	720
Dp	288531	GCAAGTCGATTTCTTGTGGCTGTGTGGCGAATACCTTGTCACAAAATGATGAGGTGGGT	288472
Qy	721	ggcgcagtcgaactgaatgaatgaanaataaactagcccaatagtttttttaacaaca	780
Dp	288471	GGCCAGTCAACTTAGTAAATGAANAATTAACATATGCCCATATGTTTATTTACCAACA	288412
Qy	781	ggaggtcatttbgcgacagtggtgcacaaatgttatactatactatgacgcaaaag	840
Dp	288411	GGAGGTCATTTTGGCGACAGATGGCTCCACCAATTTATTCATATGATGCCCAAAAGCAAAAG	288352
Qy	841	tgtgttaataatggggatgtgcacaaacgggcaacccctataagaagaagcaatgcttc	900
Dp	288351	TGCTTAATTAATGGGGATTTGCCAAACGGGCACCCCATATATAGAAAAGCATATGGCTTC	288292
Qy	901	caagtcgttcgttaaaagtttgtctatgatgaatcttgcgtgaaatacccatcaagta	960
Dp	288291	CAGCTGTTTCTGAAGATTTGGTTCATATGATGAATCTTGTCTGGAATACCCATTTACGA	288232
Qy	961	ttctaagaccacgltcaaaatlyggaatactctttaaocgataataatlygcacagta	1020
Dp	288231	TTCTAACACACCGTCAAAATVGGGAAATACTCTTTTAACGACGATATATATGCGACAGGA	288172
Qy	1021	aaaataatgcacaaactgcacaaactctctgcctcaatagatattaacaacgaacgctt	1080
Dp	288171	AAATTCATATGCCAAACATGAAACAAATTTCTCTCTAATATGATTAATAAACACGAACCGTT	288112
Qy	1081	caatgttttaatttctttataccgagacgggaagaacgtgttatacatcgtctcaggt	1140
Dp	288111	CAATTTGTTTATGTTTCTTTATATCGAAGACGCAAGAGAACTCTCTTATATCATCGCTGAGGT	288052
Qy	1141	ggtgtcaacagttatcgaccacagatgaaatlyggaanaataatttcccttaatgtgcga	1200
Dp	288051	GGTGTCAACAGTTATTCACACCAACATGAAATTAAGGAAAAATTTTCCTTTATTTGCGAA	287992
Qy	1201	ggaaaagcgcaattgatacttaccagaagaatacaatcaaggtgccttggagatataatc	1260
Dp	287991	GGAAAAGCGCAATGATATCTTACACAGAACATCAATCAAGTGCTCTGAGAGATTATATTTC	287932
Qy	1261	caaggaataatttaccgtctcgtcctgaataataacgaaacttgcgaagcgcggtcgctcat	1320
Dp	287931	CAGGAATATTTACGGTCTGCTCTGAATAATACGAAACTTGGCAAGGCGCGCGGCTTCAT	287872

QY	1321	atgagttgaagcagttacgcttactttggaagaataaagcgctggaacgcgcctgttc	1380
Db	287871	ATCACTGAAGACAGTACCGTTACTTGTGAAAGTAAACGGCGGTGGCAACGACCGCTGTCC	287812
QY	1381	aaatcggcaaaagcagcgtgcacgcttcaaacgaaggggaaaaaacaagctcgatcagc	1440
Db	287811	AAAAACGGCAAAAGGACACGCTCTACAGTTTCACCAACCAAGGGGAAAAACCAAGGCTCGATCAGC	287752
QY	1441	gtggcgacggtacagctatcttgcatacgaacgacagcaaaagcaaaaaacaagc	1500
Db	287751	GTGGCGCAGCGGTACAGTCATTTTGTGATCAGCAGGCGACACGATTAAGGCCAAAAACAAGCC	287692
QY	1501	tttagtgaatccggtctggttcagcgcgaggggtacggtgcgaactgaaatgcgataatcag	1560
Db	287691	TTTATGTGAATCCGGCTTGCTGTAGCGGCAAGGGGTACGGTGCACATGAATGCGCATTAATCAG	287632
QY	1561	ttcaaccccgcaacaacttacttctgcgcttttcgcgcgcgcgcgtttgtaattlaaacggtgcat	1620
Db	287631	TTCAACCCCGCAACAACCTAATTTTGGCGTTTTCGCGCGAGCGATTGTGATTTAAAGGGGCAT	287572
QY	1621	tcgcttctgcctccacgcgattcacaataacagatgaagggcgatgattgtcacaacacaaat	1680
Db	287571	TCGCTTGTGTTCCACCGTATTCAAAATACGATTAAGGGCGATGATTGTCAACACCAAT	287512
QY	1681	caagcaaaagaatccaccgttaccattacaggaatgaagatattgtctacaacgcgcgaat	1740
Db	287511	CAGAACAAGAAATCCACCGTTACCATTAACAGCATTAAGATATTGTCTACACCGGCAT	287452
QY	1741	aacaaagcttggaatgaacaanaaagaatgtccctacaacggttggtttgcgcgaagaagt	1800
Db	287451	AACAACACCTTGATGATACCAAAAAAGAAATTCCTACACACGTTTGTTGCGGAAAAAGAT	287392
QY	1801	acgcgcacaaagaacgcgcgcgcgtacaacttgattacacgcgcgcgcgcgaagaagacgcac	1860
Db	287391	ACGACCAAAACGAACGGGCGGCTAACCTTGTTTACCAGCCCGCGCAGAGACCGCAC	287332
QY	1861	ctgcgtctttccgcgcgcgaacaattlaacgcgcgaatcaacgaacaaacagcgcaactg	1920
Db	287331	CTGCGCTTTTCCGGCGGAAACAAATTTAAAGGCAACATCAACGGAACAAACGGCAAACTG	287272
QY	1921	tttttcagcgcgaacacacacacgcgcgcgtctacatcaattlaaacgcaacttgctgcga	1980
Db	287271	TTTTTCAGCGGCAACCAACACCGCCACCGCTACATCATTTTAAACGACCATTTGCTGCA	287212
QY	1981	aaagaagcattctctgcgcgcgcgaatactgtttggaacaaagacgtggatcaacgcgcacatt	2040
Db	287211	AAAGGCGCATTCCTCCGGGGAAATGCTGTGGACACACACTGTGATCAACCGCAATTT	287152
QY	2041	aaagcggaaactctcaaatlaaagcggaacagcggttggtttcccgcaatgttgcga	2100
Db	287151	AAAGCGGAAAACTTCACAAATTTAAAGGGGAGAGCGCGGTGGTTTCCGCAATTTTGGCA	287092
QY	2101	gtgaaagcgcatctgcatcttgagaaatcaacgcgcgaacaaatttctggtgcgaacgcat	2160
Db	287091	GTFGAAGCGCATTTGGCATTTTGAGCAATCAACGCCCAACGATTTTGTGTGCGCACCGCAT	287032
QY	2161	caagcgcaacaaactcttaacggttcgagacttggaacggttgcgaactgtgtgcga	2220
Db	287031	CAAGCGCAACAATCTGTACACGTTGAGCTGGACGGGTGTGCAAAATGTGTGCAAAA	286972
QY	2221	acaactaacgcagataaagatgatgtcttaacttgactaagaacgcagatcaacgcgcaatgtc	2280
Db	286971	ACCAATTAACCGCAATTAAGATGTGCTTCAATTAAGACCAACATCAACGGCAATGTGTC	286912
QY	2281	gattcttcgcatacagcttaataactcaagggcgttcgcaactcaacgcaactct	2340
Db	286911	GATCTGCGCGATCAGCGCTCATTTTAAATCTCACAGGGGTTGCCACACTCAACGSCATCTT	286852
QY	2341	atgtcgaatgtgcatacagttatacagtaacgccaacgcaccccaaaaacggtcaactt	2400
Db	286851	AGTCCAATATGGCGATACAGCTTTATACGTATAGCCACACACCCACCAAAACGGCAACCTT	286792

[illegible]

Db 285711	CTGCACACCCCAACCGCAGCCCAACCGCAGCGACMCCTGATGAGCGCTTATGCCATATGAC	285652
Qy 3541	ggcttgatgaaatcttcgcgcacgctcaacagcgttttcgcgctacgaagaatagac	3600
Db 285651	GGTTTGAGTGAATTTTTCGCGCACGCTCAACACAGTTTTCGCCGTACAGACGAATTAAGC	285592
Qy 3601	cgcgtatctgcgcgaagacccgcgcgaacgcgcttttgaaanaagcgcgatccgcgcgaacaa	3660
Db 285591	CGCGTATTTGCCGGAAGACCGCGCAACGCCGCTTTGGAAAGAGCGCATCCGGGACACCAAA	285532
Qy 3661	ccctccoccttcgcgaagatcttcgcgccttaacgcgcgaacaaaccgaccttcgcgaactcgt	3720
Db 285531	CACCTACCGTTGCCAAGATTTTCGCGCGCTTACCGCCACAACACCGACCTGCGCAATTCGCT	285472
Qy 3721	atgcgaanaaaccttcgcgcacgcgcgcgttcgcgcatcctgttttcgcacaaaccgcgaac	3780
Db 285471	ATGCGAATAAAACCTTCGCGACGCGCGCGCTGCGCATCTGTTTTCGCACAAACCGGACCCAA	285412
Qy 3781	aacaccttcgcgaacgcgcacgcgcgaactcgcgaacgcgtctgcgcgaacgcgcgttttcgcg	3840
Db 285411	AACACCTTTCGACGACGCGCATCGGCAACTGCGCAGCGCTTGCCACGCGCGCCCTTTTCGGG	285352
Qy 3841	caatccgcgatctgaacaggtctactactcgcgcgtcaacgcgcgcgcgcgcgcgttttagacgcgc	3900
Db 285351	CATATCGGCGATGACAGGCTGTACACTCGGCATCAACGCGCGCGGCGGCTTTTAGCAGCGGC	285292
Qy 3901	agccttccgcgaacgcgcgttcgcgaacgaacatccgcgcgcgcgtctgcgcatcaacgcgcatcag	3960
Db 285291	AACCTTTCGACGCGCATCGGACGCAAAATCCGCCCGCGGTCTGCATTACGCGCATTCAG	285232
Qy 3961	gcacgataccgcgcgcgcgttcgcgcgcatctgcgcgatctgcgcacgcgcgcacatccgcgcgcgc	4020
Db 285231	GCACGATACCGCGCGGCTTTCGCGGATTCGCGCATCGAACCGCACATCGCGGCAACGCGCC	285172
Qy 4021	tatttcgtccaaaagcgcgatccgcgtcgcgaanaagctcaatatctgcacacccgcgcgctt	4080
Db 285171	TATTTTCGTCCAAAACCGGATTTACCGCTTACGAAACGTAATATGCGCACCCCGGCTT	285112
Qy 4081	gaattcaaacgcgtaccgcgcgcgcgcattaaagagaattattcaataaacgcgcgcgaacac	4140
Db 285111	GCATTCACACCGCTACCGCGCGCGGCAATTAGGCAAGATTATTCATTCAAACCGCGCACACAC	285052
Qy 4141	atttcacatcagcgccttaattgcgcctgcgtcctatccgatgcgcgcgttcgcgcgaacgtccga	4200
Db 285051	ATTTCGATACCGCTTATTTGACCGCTCTATTACGATGCGCGCTTTCGGGCAAAATCCGA	284992
Qy 4201	acaagcgttcaataaccgcgcglatctgctcaagatcttcgcgcaaaaccgcgaagtcggaatgg	4260
Db 284991	ACACGCGTCAATATACCGCCGCTATTGGCTCAGGATTTTCGGCAAAAACCGCACGTCGGAATGG	284932
Qy 4261	ggcgtlaaacgcgcgaatacgaagaagtttcaacgcgcgtccctccacgcgtcgcgcgcgcgaagtc	4320
Db 284931	GGCGTAAACGCGGAATTCAAAGTTTCAACAGCTGTCCCTTCACGCTGCGCGCCCGCAAAAGCG	284872
Qy 4321	ccgcgaacttcgaagcgcgaacacagcgcgcgcgcgcacaaatagtgctacgcgtgcgtca 4374	
Db 284871	CCGCACACTGGAAACGCAACACACACACGCGGCGCATCAATTAGGCTTACCGCTGATTA 284818	
RESULT 5		
NME242535		
LOCUS	NME242535 4374 bp DNA linear BCT 16-JAN-2002	
DEFINITION	Neisseria meningitidis serogroup B app gene, strain B:15:PI.16.	
ACCESSION	AJ242535	
VERSION	AJ242535.2 GI:6977940	
KEYWORDS	APP gene; app protein.	
SOURCE	Neisseria meningitidis serogroup B.	
ORGANISM	Neisseria meningitidis serogroup B	
	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;	
	Neisseria.	
REFERENCE	1 (sites)	
AUTHORS	Hadji, H. A., Woolfdridge, K. G., Robinson, K. and Ala'Aldeen, D. A.	

Db 1141 GGTGTCACAGTATTATGACCCAGACTGATATATGAGAAAAATTTCTTTATTTGACGAA 1200
QY 1201 ggaagagcggaatttgatctactaccgaacatcaatgaagtgctggaagatatctc 1260
Db 1201 ggaagagcggaatttgatctactaccgaacatcaatgaagtgctggaagatatctc 1260
QY 1261 caaagagatttctagctctcgcgtgaaatacgaacttgcagagcgcgagcgctcat 1320
Db 1261 CAAGAGATTTTACGCTCTCCCTGAAATATGCAAACTTTGGCAAGCGCGCGCTTCAT 1320
QY 1321 atcagtgaaagacagctacgcttacttgaagaaatgaagcgtggaagacgagcgctctc 1380
Db 1321 ATCAGTGAAGACAGTACGCTTACTTGGAAATGAACGGCGTGGCAACGACCGCTCTCC 1380
QY 1381 aaaaatcgcaaaagcagcgtcgaagcttcaagcgaagaggaagaaacgaagctcgaacg 1440
Db 1381 AAAATCGCAAGAGCGACGCTGACGCTTCAAGCAAGGGAAGAAACCAAGGCTCATGAC 1440
QY 1441 gtgagcgagctagctacatcttctgagtcagcagcagcgaatgaagcgaagcgaagc 1500
Db 1441 GTGAGCGAGCTAGCTATTTTGGATCAGCAGGACGATGAAGGCAAAAACCAAGCC 1500
QY 1501 tttagtgaaatcgcttctgagtcagcagcgagtgagtcagctgaaatgacgaaatcag 1560
Db 1501 TTTAGTGAATCGCTTGTGTCAGCGGAGGCTACGGGTCAGCTGAATGCGGATTAATCAG 1560
QY 1561 ttcaaccccgacaaactctatcttgcgcttgcgagcgagcttctgaaatcgaagcgt 1620
Db 1561 TTCAACCCCGACAACTATTTTGGCTTTCGCGGACGCTTGGATTAAGGCGGCTAT 1620
QY 1621 tgcgtctcgttcacacgctatcaaatcgaatcgaagcgagctgagcttctgaaacgaat 1680
Db 1621 TCGCTTGTGCTTCCACGCTATTCAAATATCGCATGAAGGCGGCGGATTTGTCACACAAAT 1680
QY 1681 caaagcaaaagatccacgcttacccttaccgaagaaatgaatgctgtaaacacggaat 1740
Db 1681 CAAGACAAAGATCCACGCTTACCTTACAGCAATGAATGATTTGCTACACCGGCAAT 1740
QY 1741 aacaaacgcttggatagcaaaaagaatgtgcctacaaagcttgcgttgcgaagaaat 1800
Db 1741 AACAAACGCTTGGATAGCAAAAATAATGCTTACAAACGCTTGGTGGGCAAAAAGAT 1800
QY 1801 acgacccaaacgaacgagcgagctcaactgttctacacgagcgagcgaagacgagc 1860
Db 1801 ACGACCAAAAGCAAGGCGGCTCAACTTGTATACAGCGCGCGGCGGAGAAAGACGAC 1860
QY 1861 ctgcgtcttccgagcgagcaaatcttaagcgcaacatcagcaacaaacgagcaactg 1920
Db 1861 CTGCTGCTTCCGCGGAAACAAATTTAAACGCAACATCAACGCAAAACGGAACCTG 1920
QY 1921 tcttcagcgagcagaacaaacgagcagctacaaatctttaaagcgaacatctgtctgcaa 1980
Db 1921 TCTTTCAGCGCGAGCAACAAACCGGACGCTTAAATCAATTAAGCAACATTTGCTGCAA 1980
QY 1981 aaagagggagctcctcgcgaggaagaaatgtgtgagcaacgagcttgaacacgagctt 2040
Db 1981 AAAGAGGGATTCCTCGCGGGAATCTGTGTGGCAACGATCGATCAACCGGACATTT 2040
QY 2041 aaagcggaacactctccaaatgaagcgagcagcgagtggttcccgcaatgtgtccaa 2100
Db 2041 AAAGCGGAACATCTCAATTAAGAGCGGACAGGCGGTGTTTCCCAATGATGTCGCAAA 2100
QY 2101 gtgaagcgagcttgcgacttgcagcatcagcccaagcagcttctgtgtgcacgagct 2160
Db 2101 GTGAAGCGGATTTGCTATTTGAGCAATCAGCCCAAGCAGTTTGTGGTGCACCGCAT 2160
QY 2161 caaagccaaacacactctgacagcttgcagcagctgagcagctgagcaaatgtgtcga 2220
Db 2161 CAAAGCCACACATCTGTACAGCTTGCAGACTGAGCGGTGTCAGCAATTTGTGTGAAGAA 2220
QY 2221 accatccagcagctaaagtgatgtctcatctgactaaagcagcagcatcagcgagctgtc 2280
Db 2221 ACCATTACGAGATTAAGTGATGTCTTCAATTAAGTGAACCGGACATCAGCGGCAATGTC 2280

QY 2281 gatctgcagatcagctcatcttaactcagaggttgcacacactcgaacgagcaatctc 2340
Db 2281 GATCTGCGGATCAGCTCATCTTAATCTTCAAGGCTTGCACACACTCAACGCGCATCTT 2340
QY 2341 agtcgaacttgcagctacagcttactagctcagccaacgagccacccaacgagcaactc 2400
Db 2341 AGTCGAATGTGCGATACAGCTTATACACTAGCCACTACGCAACCCAAAGCGCAACTT 2400
QY 2401 agcctcgtgagcaatgagcccaagcaacatttgaacccaacatgaacgagcaacatc 2460
Db 2401 AGCCTGCTGAGCAATGCCCAAGCAATTTAATCAACGCAATTTAAAGCGCAACATCG 2460
QY 2461 gcttcgagcaatgtctcaatcttaactgaagcgaacgagcgttacaagagcgagctgag 2520
Db 2461 GCTTCGAGCAATGCTTCAATTAATCTTAAGCGACACCGCTTACAAAGCGGACGCTGAG 2520
QY 2521 ctctcgcgaacgctaaagcaacgtaagcgaatcgcagctcgaacgagctgagctcga 2580
Db 2521 CTCTCGGCAACGCTAAGGCAAGGTAAGCCATTCCGACACTCAACGCTAATGTCTCCTT 2580
QY 2581 gccgaatgaagcagatcttccatcttgaagcagcgagcttcccgagcaacatcagcgagc 2640
Db 2581 GCCGATTAAGCGATTTCTCATTTTGAAGCAACCGCTTTACCGGCAACATCAGCGGCG 2640
QY 2641 aaggaatcagcattacacttaaaagcagcgaatgagcagctgcgctcagcaggaatta 2700
Db 2641 AAGGATACGCGATTTACCTTAAAGACAGCAATGAGAGCGCTGCATACGCGAGAAATTA 2700
QY 2701 ggcgaatlaaaccttgaacacgcaacatlaacactcaatccgactcgcagcagctgag 2760
Db 2701 GGCATTTAAACCTTGACACAGCCACCATTTACACTCAATTCGCTTACGCAATTCGCG 2760
QY 2761 gcaagcgcaaacgagcagctgagcagcagctgagcagcagcagcagcagcagcagc 2820
Db 2761 GCAAGCGCAACCGGAGCTGCGACATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2820
QY 2821 cgttccctactatccgttaccacgagcaactcgttgaatccgcttcaacacgctgag 2880
Db 2821 CGTTCCTATTTATCTGCTTACACCGCAACTTCCGCTTAAATCCGCTTCAACGCTGAC 2880
QY 2881 gtaaacgcaaatctgaacgctcgaagcaatccgcttactgtcgaacactcgcgctac 2940
Db 2881 GTAACGCAAAATTAAGCGTCAAGGCAACATTCGCTTATGTCGGAACCTTCGCGTAC 2940
QY 2941 cgcagcgcaaatctgaagctgagcagcagctcgaagcagcttaccacttgcgctgagc 3000
Db 2941 CGCAGCGCAAAATTAAGCTGCGGAAGGTTCCGAAGGCACTTACACTTGGCGGCTAAC 3000
QY 3001 aatccgcaacgaacacttgaagcctcgaacatctgaacagctgagcagcagcagcagc 3060
Db 3001 AATACCGGCAACGAACCTTGAAGGCTTCAACAAATTTGAACGAGTGAAGGAAAGAACAC 3060
QY 3061 aaacgagcttgcgaacacacttaattcaaccttgaacaaacgaacagcagctgcgagc 3120
Db 3061 AAACGCTGTCTCGAAGCACTTAATTTCACTTCCGCAAAACGACGTCGATGCGCGGCG 3120
QY 3121 tggcgcttaccactcctcgaacagcagcagcttgcgcttgaatctcgttcaagaa 3180
Db 3121 TGGCGTTACCACTCATCTCGCAAGACGCGGAGTTCGCGCTGCAATTCGCTCAAGAA 3180
QY 3181 caaagcttctcgaacacactcgaacagcagcagcagcagcagcagcagcagcagcagc 3240
Db 3181 CAAGAGCTTTCGCAACAACTCGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 3240
QY 3241 gcgcaaacgcttgaagcagcttgcgagcagcagcagcagcagcagcagcagcagcagc 3300
Db 3241 GCGCAAGGCTTGAAGCGCTGATTTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3300
QY 3301 gtgcgcaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3360
Db 3301 GTTGCGCAACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3360

[illegible]

OY	1141	gggtcacaagctatcgcaccagcagctgaaabaatygaaataattccttatttgacga	1200
Db	1141	GGTGTCAACAGTTATCGACCCACACATGGAATATGGAAGAAAATATTTCTTTATTTGACGA	1200
OY	1201	ggaaaaagcgaattgatacttaccagcaacatcaatcaagtgctgcgagagatctatctc	1260
Db	1201	GGAAAGAGCGAATTGATCTTACCAATCAATCAACCGGAGGCTGGAGAGTTATATTC	1260
OY	1261	caaggaatttaccggtctgcctgaanaataacgaagacttgcgaagcgcgaggttcat	1320
Db	1261	CAAGAGCAATTTTACGGTCTGCTGTGAACACAGAACCTGGCAAGCGCGGCGTTTAT	1320
OY	1321	atcagtgaaagacagtaacggttacttggaaaglaaacgycgtlygcaacagccgcgtcc	1380
Db	1321	ATCAGTGAAGACATNATCCGTTACTTGGAAAGTCATATGCGCTGGCAAGCGCGCTGCC	1380
OY	1381	aaatcgcgcaaaagcagcgcgcgacttcaagccaaagggaaaccaaagctctgcatacg	1440
Db	1381	AAATTCGGCAAAAGCGACGCTGCACGTTCAACGCCAAAGGGGAACCAAGGCTCATAGC	1440
OY	1441	gtgagcgacaggtacagtaacttlygatcagcagcgacgacgataaagcacaanaaacagcc	1500
Db	1441	GTGGCGCACCGTTAAAGTTATTTAGATTAACAACGACGATGAAATATATPAAAAACAAACC	1500
OY	1501	tttagtgaaatcgcgctctgycgcagcggcgaggggtacgylbcaactbaatgcgataatcag	1560
Db	1501	TTTAGTGAATCGCGCTTGGTCACCGGAGGGGTACGCGTCAACTGAATGCGATATACG	1560
OY	1561	ttcagcccccbaaacactatttcgcgcttcgsgggsgagcgttttgattttaaocggagat	1620
Db	1561	TTCAACCCCCGACAAACCTATTTTCGGCTTTGCGGGGACGTTTGSATTTGACACGGCAT	1620
OY	1621	tcgcttcgtctccacgctatctacaabaatbcgatalgaagggcgatgtgcacaccaat	1680
Db	1621	TCGCTTTCTGCTCCACCGTATTTCAAAATACGAGTAAGAGGGCGCATATGTCAACCCAT	1680
OY	1681	caagacaagaatcccaacgcttaccatctacagagaaataagataltgctacaaocggcat	1740
Db	1681	CAGACAAAGATTCACCGTTCACATTACAGCGAATAAAGATTTGCTACAAACGGCGCAT	1740
OY	1741	aacaaacgcttggatagcaaaaaaagaatgcctacacagcttggcttgcgagaaagat	1800
Db	1741	AACAACACCTTGGATATACCAAAAAAGAAATTCGCTCAACAGTTGGTTGGCCAGAAAT	1800
OY	1801	acgacccaaacgaacggcgcgctccaacactgttctacagcccgcgacgaagacccgccc	1860
Db	1801	ACGACCAAAAGSAGCGGGCGGCTCAACCTGTTTACCAACCGCGCTGCAAGAACCGCAC	1860
OY	1861	ctgcgtctctccgcgcggaacaatttaaaccgacacatcaacgcaaaaacaaagcgcaactg	1920
Db	1861	CTGCTGCTTCTCCGGCGGAGAACAAATTTAAAGGCAACATACGCCAANAACCGCAACTG	1920
OY	1921	tttttcagcgcgaacaacaacacccgcgcctacaaatcatttaaagacatctgtgtcgcaaa	1980
Db	1921	TTTTTTCAGCGGCAACCAACACCGGCAACGCGCTCAATCATTTTAAAGACCATTTGTCGCA	1980
OY	1981	aaagaagcattctctgcgcggggaatctcgttggagaaagacatctgataaacccgcattt	2040
Db	1981	AAAGAGGCGCATTTCTCCCGGGGAATCTGTGTGGCAACAGCTGGATCAACGGCACTTT	2040
OY	2041	aaagcggaaaccttccaaatlaaagcgagacagcggtgttctccgcaatgtgttcgcaaa	2100
Db	2041	AAAGCGGAAACTTCCAAATTTAAAGGGCGGACAGCGGTGGTTCCCGCAATCTTCCCAAA	2100
OY	2101	gtlgaagcgagtttgcaatttgagcaatcaagcccaagagatttttgggtgtgcacgcgat	2160
Db	2101	GTGAAAGCGCATTTGCAATTTGAGCAATCAAGCCCAAGCATTTTGTGTGTGCAACCGCAT	2160
OY	2161	caaagccacacatctgtacagttcgcgactgtgacggtgtctgacaaatctgtgtcgaaaaa	2220
Db	2161	CAAAAGCCACAAATCTGTACACGTTGGAGATGGACGGGTCTGACAAATTTGTGTGCAAAAA	2220

QY 2221 accattaccgaagataaagtgatgtctcatatgaactaagacgacatcagcggaatgtc 2280
|||||
Db 2221 accattaccgaagataaagtgatgtctcatatgaactaagacgacatcagcggaatgtc 2280
QY 2281 gatctggcgaatgaagctatattaaatctcaagagcttgcaacatcaagcaatctc 2340
|||||
Db 2281 gatctggcgaatgaagctatattaaatctcaagagcttgcaacatcaagcaatctc 2340
QY 2341 agtgcgaatggcgatacgaattatagtcagacacacgacacccaagaagcaactc 2400
|||||
Db 2341 agtgcgaatggcgatacgaattatagtcagacacacgacacccaagaagcaactc 2400
QY 2401 agcctgtgggaatggcccaagcaacatttaactcaagcacaatcaagcgcaacatc 2460
|||||
Db 2401 agcctgtgggaatggcccaagcaacatttaactcaagcacaatcaagcgcaacatc 2460
QY 2461 gcttcgggaatggctcatatcatatgaagcagcagcagcagcagcagcagcagc 2520
|||||
Db 2461 gcttcgggaatggctcatatcatatgaagcagcagcagcagcagcagcagcagc 2520
QY 2521 ctctccggaacgctaaagcaacgtaagccatctccgcaatcaagcgttaatgtctcccta 2580
|||||
Db 2521 ctctccggaacgctaaagcaacgtaagccatctccgcaatcaagcgttaatgtctcccta 2580
QY 2581 gccgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagc 2640
|||||
Db 2581 gccgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagc 2640
QY 2641 aaggaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagc 2700
|||||
Db 2641 aaggaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagc 2700
QY 2701 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagc 2760
|||||
Db 2701 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagc 2760
QY 2761 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagc 2820
|||||
Db 2761 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagc 2820
QY 2821 gcttcctatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 2880
|||||
Db 2821 gcttcctatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 2880
QY 2881 gtaaacggaacaaatgaagcagcagcagcagcagcagcagcagcagcagcagcagc 2940
|||||
Db 2881 gtaaacggaacaaatgaagcagcagcagcagcagcagcagcagcagcagcagcagc 2940
QY 2941 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagc 3000
|||||
Db 2941 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagc 3000
QY 2992 aatgacggaacaaatgaagcagcagcagcagcagcagcagcagcagcagcagcagc 3060
|||||
Db 2992 aatgacggaacaaatgaagcagcagcagcagcagcagcagcagcagcagcagcagc 3060
QY 3061 aatgacggaacaaatgaagcagcagcagcagcagcagcagcagcagcagcagcagc 3120
|||||
Db 3061 aatgacggaacaaatgaagcagcagcagcagcagcagcagcagcagcagcagcagc 3120
QY 3121 tggcgtttacaaatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 3180
|||||
Db 3121 tggcgtttacaaatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 3180
QY 3181 caaagagcttccgcaacaaatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 3240
|||||
Db 3181 caaagagcttccgcaacaaatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 3240
QY 3241 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagc 3300
|||||
Db 3241 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagc 3300
QY 3301 gttgcgaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3360
|||||

Db 3292 gttgcgaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3351
|||||
QY 3361 gagaacaaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3420
|||||
Db 3361 gagaacaaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3420
QY 3352 gagaacaaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3411
|||||
Db 3352 gagaacaaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3411
QY 3421 accgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3480
|||||
Db 3421 accgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3480
QY 3481 ctgcaaccccaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3540
|||||
Db 3481 ctgcaaccccaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3540
QY 3472 ctgcaaccccaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3531
|||||
Db 3472 ctgcaaccccaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3531
QY 3541 gttgcgtttacaaatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 3600
|||||
Db 3541 gttgcgtttacaaatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 3600
QY 3552 gttgcgtttacaaatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 3660
|||||
Db 3552 gttgcgtttacaaatcatcatcatcatcatcatcatcatcatcatcatcatcatcatcat 3660
QY 3601 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 3660
|||||
Db 3601 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 3660
QY 3661 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 3720
|||||
Db 3661 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 3720
QY 3652 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 3711
|||||
Db 3652 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 3711
QY 3721 atgcaacaaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3780
|||||
Db 3721 atgcaacaaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3780
QY 3781 aacacccctgaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3840
|||||
Db 3781 aacacccctgaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3840
QY 3772 aacacccctgaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3831
|||||
Db 3772 aacacccctgaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3831
QY 3841 caataacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3900
|||||
Db 3841 caataacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3900
QY 3832 caataacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3891
|||||
Db 3832 caataacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3891
QY 3901 agccttcgaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3960
|||||
Db 3901 agccttcgaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3960
QY 3961 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 4020
|||||
Db 3961 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 4020
QY 3952 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 4011
|||||
Db 3952 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 4011
QY 4021 tatttcgttcgaacaaatgaagcagcagcagcagcagcagcagcagcagcagcagcagc 4080
|||||
Db 4021 tatttcgttcgaacaaatgaagcagcagcagcagcagcagcagcagcagcagcagcagc 4080
QY 4081 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 4140
|||||
Db 4081 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 4140
QY 4072 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 4131
|||||
Db 4072 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 4131
QY 4141 atttcgaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4200
|||||
Db 4141 atttcgaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4200
QY 4132 atttcgaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4191
|||||
Db 4132 atttcgaacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4191
QY 4201 aacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4260
|||||
Db 4201 aacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 4260
QY 4261 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 4320
|||||
Db 4261 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 4320
QY 4252 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 4311
|||||
Db 4252 ggcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagc 4311
QY 4321 cgcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagcagc 4374
|||||
Db 4321 cgcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagcagc 4374
QY 4312 cgcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagcagc 4365
|||||
Db 4312 cgcgaatgaagcagatcatcatatgaagcagcagcagcagcagcagcagcagcagcagcagc 4365

RESULT 7


```

|||||
Db 3952 AGAGCAAAATCCCGCGCGTGTGATTTAGGCAATTAGGCAAGTACCGCCAGCT 4011
|||
Qy 3979 ttcggcgatctcgcaatcgcaacgcgacgagcgacgcatcttcgtccaaagcg 4038
|||||
Db 4012 TTCGCGGATTCGCAATCGAACCCGACATCGGCGCAACGCCCTATTTGTCACAAAACG 4071
|||
Qy 4039 gattacgcctaaagcaatcacaatcgcaaccccgcccttcgacatcaacgcgacgc 4098
|||||
Db 4072 GATTACCGGATTCGCAAAACGCAATATGCGCAACCCCGGCCCTTGCAATACCGCTACCGC 4131
|||
Qy 4099 gcgagcatlaagcagatattcatcaaacgcgcgcgaacacattccatcaacgcctat 4158
|||||
Db 4132 GCGGCGCATTAAGGAGATTTGATTTCAATTAACCGCGCAACATTTTCATCACGCCCTTAT 4191
|||
Qy 4159 ttgagcctctcattacagatcgccgcttcggaacagtcgcgaacgcgcatcaatcgcc 4218
|||||
Db 4192 TTGAGCGCTCTTATACCGGATTCGCGGCAAAAGTCCGACGCCGCTCATTTACGCCGC 4251
|||
Qy 4219 gttatgctcaggaatttcgcaaaaacccgcagtcgcgaatgagtgagtgtaacgcgaatc 4278
|||||
Db 4252 GTATTGGCGGAGGATTCGCGCAAAACCGCAGTGGGATGGGGGTAACCGCCGAAATC 4311
|||
Qy 4279 aaaggttcaagctctccctcacaagctcgccgcgcgaacgcgcgaactggaagcgca 4338
|||||
Db 4312 AAAGGTTTCAAGCTGTGCTCCCTCCAGCTGCCGCCGCAAGGGGCCCATTTGGAACGCGAG 4371
|||
Qy 4339 cacagcgcgagcatcaatattagctaccgctgta 4374
|||||
Db 4372 CACAGCGCGGATTCAAATTAAGGCTACCGCTGTGA 4407
|||||

RESULT 8
LOCUS NME296276 4350 bp DNA linear BCT 14-JAN-2002
DEFINITION Neisseria meningitidis app gene for adhesion penetration protein,
ACCESSION NME296276
VERSION A0296276.1 GI:11071862
KEYWORDS adhesion penetration protein; App gene.
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE
1 (Sites)
Hadi, H.A., Wooldridge, K.G., Robinson, K. and Ala'Aldeen, D.A.
Identification and characterization of App: an immunogenic
autoantigen protein of Neisseria meningitidis
Mol. Microbiol. 41 (3), 611-623 (2001)
2 (bases 1 to 4350)
Ala'Aldeen, D.A.
Direct Submission
Submitted (27-OCT-2000) Ala'Aldeen D.A., Microbiology, University
of Nottingham, Molecular Bacteriology and Immunology Group,
University Hospital, Nottingham NG7 2UH, UNITED KINGDOM
LOCATION/Qualifiers
1. 4350
/organism="Neisseria meningitidis"
/strain="23515"
/db_xref="taxon:487"
/country="Saudi Arabia"
/note="serogroup A"
1. 4350
/gene="app"
1. 4350
/gene="app"
/gene="app"
/codon_start=1
/transl_table=1
/product="adhesion penetration protein"
/protein_id="CAC14670.1"
/db_xref="GI:11071863"
/db_xref="SPTREMBL:Q9F325"

```

```

/translation="MKTDRKTEETHRKAPKTRGIRSPAYLAICSLILPOAGH
TYRGINVOYEDFAENKGGKFAVGAKDIEVYKKGELGKSMTRKAPMIDFSVSNAGV
ALVGDQITVSAHNHCYNNNDPQAGEGRPDQDRPSYOTYKNNKPKPNHSNGDYH
PRLHKFTYDAEPVTHSDMKRNTISDKRIFERYRIGSHHTWKIDDKHEDLSGCA
WLTGNTHMGMGNNGVSLISGDVRAHNDGPMPIAGAGSGSSPMTLYDTNKKMLL
NGVLQOTYPPVSGRNGFOLIRKDMFVYDLYKGDHTYFEEFRSNGHSFSSNNGTG
VTENNERVSNPKLEVOYVRLDESLINEDYKDEPYRAAGVNOYRRLNNGENLSFIDG
NGKLTLSNNINOGAGGLYFEGDFTVSENNETMGAGVHISEDSITVWKVGNVNDRL
SKIGKTLHYOAKGENOGSISVGDGYTIDQADKKKAKAFSEGTGVLGAGGYOLNA
DNQFNPKLYKCFRGRGLDNGHSLHSTIONDEGMIVHNNHTTSTVITGNESTI
TOPSCKNINRLNPKETAYNKNPEKPTTNGNLNLYOPAADRLILLSTGNNLNG
NITQNGRLPESKPTPHATNHLDSGSKMGITIQGETWPNQDINTREKRNHTIG
GOAVISRNVAAVEGDWHLNNAQAVFVGAPNLSITCTRSMDPTLCNCEVTITDDYK
IASLTKTDISGNVADHNLGLTGLNLSLNDPDTVTYVSHNTONNLSYCN
AOAFENQATLNGNTSAGNASFNLSNNAAGNSLTLSNNAKANYSHSALNGNVSJADK
AAPHENSRFTGOLSGSKDPTALHLKDSNMTLPSTGELGNLDNATITLNSAYRHDAK
GAOTGSVDPTRRSRSLSVPTPSVSEPHNTLVNKGKLNFGGTFRPMSLELYRS
DKTLAESSSECTVTLAVNNGNEPVSIDOLTVVEGKDKPKLSENLNFTLIONEHDAG
MRVOLIREDGFRRLNPKYKBEISDKTLKAKAKOAREDNOSIDALIAGRDAERT
ESVAPRQAGENVGIMQAEERKRVQADKOSLAKOREKETRPATTAFRARARR
DLDPQPOPOPOPOHDLISRYANSGLSEFSATLNSVFAVODELDREYFAEDRRNAV
TSGLRDKTRHRSODEFRAYROQTDRLQJOMKNLSGRVILFSHNRENTFDDIGNS
ARLHAGAVFGQYIGIRGIDIGISTAGFSGSLSDGIGKIRRYVHGIQARYRAGG
GFGEIPIYIGATRYVOKADRYENVNINATPGLAFNRYRAGIKADYSEFPADHISITPY
LSLSTYDASGKVRPRTVAVLADQFCKTRSAEMGVNAIEKFTLSLHAANAAGKPOLE
AQSAGIKLIGYRM"
BASE COUNT 1236 a 1185 c 1073 g 856 t
ORIGIN
Query Match 81.5%; Score 3566.8; DB 1; Length 4350;
Best Local Similarity 89.7%; Pred. No. 0;
Matches 3936; Conservative 0; Mismatches 402; Indels 48; Gaps 8;

```


[illegible][illegible]

[illegible][illegible]

[illegible]

D	b	181	AAAGCAAGTTTGCAGTCGGGGGGAAGATATTGAGCTMTACACAAAAGGAGGATTG	240
O	y	241	gtcggcaaatcaatgacaaaagcccgatgattgatttctcgtgtgtcgtgtaagc	300
D	b	241	GTCCGCAAAATCAAGACAAAAGCCCGATGATTGATTTTCTGTGTGTCGGTAAACGGC	300
O	y	301	gtcggcagcatcgtgtggtcgatcaatatctgtgaagctgtgcacataaagcgtcatlac	360
D	b	301	GTGGCGGCAATTGGTGGGCGATTCAAATATTGTGAGGTGGACATTAACGGCGGTATTAAC	360
O	y	361	aacgttgatttctgtgtcgtgaagaagaatcccgatcaacacgtcttactataaatt	420
D	b	361	AACGTTGATTTTGGTGGGGAAGAAATCCGATACACCGTTTCTTACCAAAAT	420
O	y	421	gtgaacggaataataaagcagggaacccaagccatcccttaatgagcgttatcat	480
D	b	421	GTGAAGAAGATATTATTAAAGCCGACAAATCA---CACCTTACAAAGCGGATTAACAT	477
O	y	481	atgcgcgtctgtcaataattgtcagagatgacgaacccgtttgaaatgacccgttata	540
D	b	478	ATGCGCGCTTTGCAATTAATTTGTACAGATGACGAACCTGTGAAATGACGATCAATG	537
O	y	541	gatgycggaataatcatalcgatacaaatatcaccgcagccgtgtcgtatctgttggcagc	600
D	b	538	AGGGGAATACCTATTTCGCGATTAAGAAATAATCCGAGGTGTGCGCATCGGCTCAGSA	597
O	y	601	aggcacattctggcatctgatatgaagaatgagcccaataacccggagaagtcatataatt	660
D	b	598	CACCACTATTATGGCTTATGATGATGA---CAAACGGCGGATTTATCTTACATCCGCGC	651
O	y	661	gcaagctgcattctctgtgtcgtgtggtgcacataccttgcacaaaatgacaggtgt	720
D	b	652	GCA-----TGTTAATTGGCGCAATACACATATGACGAGGTGGGAATTAAT	699
O	y	721	ggcaacgtcaactagtagtgaaaaaataaacaataagcccatagtgttcttaccaca	780
D	b	700	GCGGTATTATTGTGACGGGGAGTGCAGGATCCGCAATGCCAACAGACTATTGCGCTTATGCCGATT	759
O	y	781	ggaagctcaatttggcgcagctgtgctcacaagtattatcattatgatactgccccaaagcaaaag	840
D	b	760	GCAGGTGGCGACGCGACGAGCGGTCCGCAAGTATTATTATATCMAAACAATATAA	819
O	y	841	tgttaataatggtggtatgacaaagcgggaaccccatatgaaagaaagcaatgtctc	900
D	b	820	TGGCGCTCAACGGAGTTTACAAACCGCTACCCCTTATTCCGACGGGAAACGGTTTC	879
O	y	901	cagctggtctgtaagaatgtgtctatgatgaataccttctgtgagatacccatcagta	960
D	b	880	CAGGTGATACCAAAAGATTGGTTCTACGATGACATTTCAGAGCGCATACATATCCGTC	939
O	y	961	tctctgcagcaacgttcaaaatggaatatctctctttaaagcagataataatggaacaga	1020
D	b	940	TMTTTTGAACCGCCGACGTAACGACATTTTCTCTTATACCCAAACAACAGCTACGGGT	999
O	y	1021	aaatcaatgccaaacatgaaacacaattctctgctaataagatlaaaacaacgaacgctt	1080
D	b	1000	ACGTTAACGAACACACGAAGAAAGTTCMAATCAAA---GCTTAAATACAGACAGTC	1056
O	y	1081	caattgttctaattgttctctctcgcgaacgaagaagaaccttcttatatcgtcaggt	1140
D	b	1057	CGACGTGTTGACGATCTTTTAATGAACACGTATTAAGAAACCACTTT---ACCGGGAGGG	1113
O	y	1141	ggtgtcaaacagttatcgaccacgaactgtaataatgtagaanaataattctccttatgacga	1200
D	b	1114	GGTGTTAATTCAGTACCGCTCAAGGTTTAAACACGGTCAAAAACCTTTCTTTTTCGATTTC	1177
O	y	1201	ggaanaagcgaattgatatacttaccagaacaacatcaatgaagtgctgtggaattatatactc	1260
D	b	1174	GGCAACGGCAACATCATCTTATCAAAACATCATCAACCAAGCGCGGGCTTTGTATTTT	1233
O	y	1261	caaggagaatttaaggtctgcctcgtgaataataagaaatcttggaaagcgcggcgcttcat	1320

Dh 1234 GAAGGTGATTTTACGGTCTCGCCTGAAACAAAGAACGTGGCAAGCGCGGCGCTTCAT 1293
Qy 1321 atcaatgaagacagtaacgtaacttgaagaagtaaacgagcgttgaagaagcgcctgttc 1380
Db 1294 ATCACTGAAGACAGTACGCTTACTTGGAAAGTAAAGCGGCTGGCAACGACCGCTGTCC 1353
Qy 1381 aaatcgcgaagaagcagcgtcagcgttcaagcgaagaaggaaggaagcgtcgtacgc 1440
Db 1354 AAAATCGGCAAGGAGGACGCTGACGCTTCAAGCAAGGGAAGGAAACCAAGGCTCGATCAAC 1413
Qy 1441 gtggcgacggtacagtaacttcttgatcagcagcagcagcagcagcagcagcagcagc 1500
Db 1414 GGGGCGAGGAGTACGATCTTTTGGATCAGCGAGAGAGATTAAGCGAAAAACAAGCC 1473
Qy 1501 tttagtgaatcgtcgttgcagcagcagcagcagcagcagcagcagcagcagcagcagc 1560
Db 1474 TTTACTGAATTCGCTTGTGTCAGCGGCGAGGGGTACGCTCACTGATGATCCGATATCAG 1533
Qy 1561 ttcaaccccgacaacatctatcttgcgcttcgagcagcagcagcagcagcagcagcagc 1620
Db 1534 TTCAACCCCGACAACCTCTATTTCGGCTTTCCGCGGAGACGTTTGGATTTAAACGGGCAT 1593
Qy 1621 tgcgttctgcgttcacgtaataacacagcagcagcagcagcagcagcagcagcagcagc 1680
Db 1594 TCGCTTCTGCTCCACCGTATTCAAAATACCGATGAAGGGCGATGATGNCNATATAT 1653
Qy 1681 caagcagaagaatccacgcttaccatcagcagcagcagcagcagcagcagcagcagcagc 1737
Db 1654 GGCACACACATCCACCGCTTACCTTACAGGAGATGAAGAGTAAAGTAAACACCGATGCT 1713
Qy 1738 ---aatacaacagcttggatagcaaaaagaatctgctcacaacgcttggcttgcagag 1794
Db 1714 AAGATATATCAATAGCTAATTTATACGAAAGAAATTTGCTTACACAGGTTGGTTGGCGAG 1773
Qy 1795 aaagaacagcagaagaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1854
Db 1774 AAAGTACGACCAAAACGACGCGGCTCAACCTTGTTCACGACCGCGCCAGAGAC 1833
Qy 1855 cgcacccgtcgtcttcgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1914
Db 1834 CGCACCCGCTGCTTTCGCGGGAACAAATTTAAACGCGACATCCGCAACCAACAGCG 1893
Qy 1915 aaactgttttcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1974
Db 1894 AAACGTTTTTCAGCGGAGACGACGACCGCTTCAATCTTTAGGAAGCGGCTG 1953
Qy 1975 tgcgaagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2034
Db 1954 TCAAAAATGGAAGGTATCCCAAGAGGAAATCGTGTGGACAAAGACTGATCNAACGCG 2013
Qy 2035 acatttaagcgaagaacttccaatlaaaggcgaagcagcagcagcagcagcagcagcagc 2094
Db 2014 ACGTTTAAAGGGGAAAAATTTCCATTTTACGGCGGCGAGCGGATTTCCCGCATGTT 2073
Qy 2095 gccaagtgaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2154
Db 2074 GCCAAGGAGGAGCGATTCGATTTGAGCAATCACGCCCAAGCGATTTGGTGTCCGA 2133
Qy 2155 ccgcatcaaaagcacaacatctgtacacgttcgagcagcagcagcagcagcagcagcagc 2214
Db 2134 CCGCATCAAAACCTTACATCTGTACACGTTTCGACTGACNGCTCTACAAATTTGTGTC 2193
Qy 2215 gaagaacacatccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2274
Db 2194 GAANAANACATTTACCGAGATTAAGTGTCTTCAATGACTAAGACNGACNTNAGCGCG 2253
Qy 2275 aatgcagcttcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2334
Db 2254 ANTGTNAGNCTNNCCNATNAGGNTNNTNNAANCTCNCNGGCGNTGNCNACTNNAAGGC 2313
Qy 2335 aatcttagtcgaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2394
Db 2314 AATCTTAGTGCAATGCGGATTCAGGTTATACAGTCAAGCCCAACGCGCAACCAAAACGCG 2373

Qy 2395 aaacttagcgttggcgaatgcccgaagcaacatcttaacgaagcagcagcagcagcagc 2454
Db 2374 AACCTTAGCTCTGTGGCAATGCCCCAAGCAATTTATCAAGCCCATTTAAACGGCAC 2433
Qy 2455 acatcgtcttcggaatgcttcaatlaaagcagcagcagcagcagcagcagcagcagcagc 2514
Db 2434 NCATCGGNTTCGGGCAATGCTTCATTTAATCTAAGCAACACCGCGCAACAAACGGCAGT 2493
Qy 2515 ctgacgcttcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2574
Db 2494 CTGACGCTTTCGACACGCTTAAGGCAACGTAAAGCATTCGCGACTCAACGGCAATGTC 2553
Qy 2575 tccctagccgaataagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2634
Db 2554 TCCCTAGCCGATTAAGGCAATTTTCCATTTTGAACACCGCGCTTACCGGCAACCTACAC 2613
Qy 2635 ggcgcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2694
Db 2614 GCGACGAGGAGACGATTTACCTTAAGAGACAGGAGTGAAGCGTGCCTGACGAGCAG 2673
Qy 2695 gaattagcaatlaaaccttgacaagcagcagcagcagcagcagcagcagcagcagcagc 2754
Db 2674 GAATTAGGCAATTTAACTTGAACAGCGCACCATTTACCTCAATTCGCGCTATCGCCAC 2733
Qy 2755 gatgcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2814
Db 2734 GATGCTGCAGGCGCGCAACCGCGAGNGTGTACACCGCGCGCGCGGCTTGGCGCT 2793
Qy 2815 tgcgagcgttccctatctacgttaccagcagcagcagcagcagcagcagcagcagcagc 2874
Db 2794 -----TCCCTATTTATCCCTTACCGCGCACCTTGGTGAATCCGTTTCAACAG 2844
Qy 2875 ctgacggttaaacgcaaatctgaagcagcagcagcagcagcagcagcagcagcagcagcagc 2934
Db 2845 CTGACGCTTAAGCGCAATTTGAACGTCAGAGACATTCGCTTATGTGCGAATCTTC 2904
Qy 2935 ggtaccgagcagcaaatctgaagcagcagcagcagcagcagcagcagcagcagcagcagc 2994
Db 2905 GCGTACCCAGAGGACAAATTTGAGCTGCGGGAATTTCCGAAGNACTTACACTTGGCG 2964
Qy 2995 gtcaacaataccggaagaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3054
Db 2965 GTCAACATACCGGCAAGACACCTTAAGCTCTGATCAATTTAGCGTGTGAAGGAGAA 3024
Qy 3055 gacacaaacgcgttccgaagaacatcttaacccctgcaaaacgaaacagcagcagcagc 3114
Db 3025 GACAACAAACCGCTGTCCGAAACCTTAATTTCACTGCAAAACGAAACGACGTCGATGCC 3084
Qy 3115 ggcgagtgagcgttaccacatccggaagaagcagcagcagcagcagcagcagcagcagc 3174
Db 3085 GCGCGTGGCTTACCACTATCCGCAAGAGCGGAGGCTTCCGCTCATATATCCGCTC 3144
Qy 3175 aaagaacaagaagccttcgcaaaaacatctgcaagcagcagcagcagcagcagcagcagc 3234
Db 3145 AAAGAACAAAGAGCTTTCGACAAACCTCGCAAGGAGGAGGCAAAACAGGCGGAAAA 3204
Qy 3225 gacacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3294
Db 3205 GACACGCGCAAGGCTTACGCGCTGATGTGGCGGCGGAGGAGGCGCGGCAAGAGCA 3264
Qy 3295 gaaagcgttcgcaaacgc 3354
Db 3265 GAAAGCGTTCGCAACCGCGCGCGCGCGCGCGGCAAGGAGGAGGAGGAGGAGGAGGAGG 3324
Qy 3355 gaggaagaagaagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3414
Db 3325 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3384
Qy 3415 ggcgaacccgc 3474
Db 3385 GCGGAAACCGCGCGGNTACACGCGCTTCCCGCGCGCGCGCGCGCGCGCGGAGATTGG 3444

D	b	691	NNNNNNNNNNNNNNNNNNNNNNNNNNNNGGCTCACCAGTGTATCTATGATGCCCAAGA	750
O	y	887	aagtggttaattatgatgggatatgcgaacggcgcaacccttatataygaaagaacaatgy	896
D	b	751	AAAGGGTTAATTATGATGGGTATTCGAACAGGGCAACCCCTATATGAAAAAGCAATGG	810
O	y	897	ctccagcctggttcgtlaaagaattggttctctaigaatactctgtcgtgagataccactc	956
D	b	811	CTTCACACTGGTGTAATAATGGTGTCTATGATGAATCTTGTGCAGATACCATTC	870
O	y	957	agtcttctacgaaccagcgttaaataatgggaatactctttaacgcgataataatgagac	1016
D	b	871	AGTATTTCTACGACACCGGTCAAAATGGGAATATCTTTTAACGAGATAATATAGGAC	930
O	y	1017	aggaaaaaatcaatgccaacaatgaaccaaatctctcgtctaatagaatlaaacaacgcagc	1076
D	b	931	AGGAAAATCATATCCCAACATGACAACATTCCTGCTTATAGATATATAAACACGAC	990
O	y	1077	cgtcataatggttaatgattctcttacccagaagaagaagaacctggttatcatctgcgc	1136
D	b	991	C GTTCATTTGTTAATGTTTTCTTTATCCGACGACGAAAGAACCTGTTATCATCTCTC	1050
O	y	1137	aggtgtygtcaacagttatcgaccacgactgaaataygaanaabatlctcttatga	1196
D	b	1051	AGGGGTGTCAACAGTTATGACCCGACGACGATATATGAGAAATATTTCTTATTTGA	1110
O	y	1197	cgaaggaagaagcgaatttga tacttaacgaaccaatcaactcaagtgctgagagatta	1256
D	b	1111	CGAAGAAAAGGCAATTGATGACTTACGACGACATCATCAAGGTCGTGSAGATTTTA	1170
O	y	1257	tttccaaggagatttctcgctcgcgtgaataa laacgnaacttgcgaacgcgcgcgt	1316
D	b	1171	TTTCCAAGGAAATTTTAGCGTCTCGCTGAAATAACGAACTTGCAAGGCGCGGCT	1230
O	y	1317	tcatatcagtgaagcagctaac tcttgaagaatgaacgcgcgtgcgaacgcagccct	1376
D	b	1231	T CATATCATGTCAGACAGTACCGTACTTGGAAGTAACGGCGTGCAAGACGCGCT	1290
O	y	1377	gtccaaatcgcgcaaaagcgcctgtcacgtltaagcgaagggaacccaagctcat	1438
D	b	1291	G TCCAAAATCGCGAAAGCGCACGCTG --- NNNNNNNNNNNNNNNNNNNNNNNN	1347
O	y	1437	caagcgtggcgacgctacgltcatlttgtagtcagcagcagacgaataaagcaaaaaa	1496
D	b	1348	NN	1407
O	y	1497	agccttagtgaatcgcgttgtcagcgcagggtagcgtcaactgaatgcgtata	1556
D	b	1408	NN	1467
O	y	1557	tcagttcaaccocgcgaactcatattcgcttcgcgcgcgcgacttgattataacgcg	1616
D	b	1468	NN	1527
O	y	1617	gcattcgcttcgttccacogtaltcaaaaacoga tgaaggcgataltgttaacca	1676
D	b	1528	NN	1587
O	y	1677	caatcagacaagaatccacgcttacocatlaagcgaataagalattgtcaacaacg	1736
D	b	1568	NN	1647
O	y	1737	caatacaacagcttgatagcaaaaaagaalttgcttaacaacgtltgittgcyagaa	1796
D	b	1648	NN	1707
O	y	1797	agatagaccaaaacgaagcgcgtcaacctlyttacagccgcgcgcgaagaccg	1856
D	b	1708	NN	1767
O	y	1857	caaccctgccttccgcgcgaaacaattaaacgcgaatccacatccacaaacaacgcgaa	1916

[illegible]

[illegible]

OY	4007	tcggcgcaacgagcgcatttcgctccaaagaagagattacgcctacgaaacgcgaatcgc	4066
Db	3988	tcggcgcaacgagcgcatttcgctccaaagaagagattacgcctacgaaacgcgaatcgc	4047
OY	4067	ccaccccgagccttgatcacaacgcgtacgcgcggggcatctaaagcagattcttcatca	4126
Db	4048	ccaccccgagccttgatcacaacgcgtacgcgcggggcatctaaagcagattcttcatca	4107
OY	4127	aacgcgcgaacacattccatcacgccttattgacctgtccctatacgaatgcgcgt	4186
Db	4108	aacgcgcgcgaacacattccatcacgccttattgacctgtccctatacgaatgcgcgt	4167
OY	4187	cgggcaagaatccgaaacacgcgtcacaatccgcgctatttgcttaagatttcggcaaaccc	4246
Db	4168	cgggcaagaatccgaaacacgcgtcacaatccgcgctatttgcttaagatttcggcaaaccc	4227
OY	4247	gcagtcgcgaatcggggcgctaaacgcgcgaatacaaaagtttcacgcgtcctccacgctg	4306
Db	4228	gcagtcgcgaatcggggcgctaaacgcgcgaatacaaaagtttcacgcgtcctccacgctg	4287
OY	4307	cgcgccgcgaagaagccgcgcacgtcgggaagcgaacacacacgcgcgcgagatcaattagctacc	4366
Db	4288	cgcgccgcgcgaagaagccgcgcacgtcgggaagcgaacacacacgcgcgcgagatcaattagctacc	4347
OY	4367	gcctggttaa 4374	
Db	4348	gctgtgttaa 4355	
RESULT	12		
LOCUS	U32710/c	13818 bp	NMA linear BCT 29-MAY-1998
DEFINITION	Haemophilus influenzae Rd section 25 of 163 of the complete genome		
ACCESSION	U32710 L42023		
VERSION	U32710.1 GI:1573200		
KEYWORDS			
SOURCE	Haemophilus influenzae Rd.		
ORGANISM	Haemophilus influenzae Rd		
	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;		
	Haemophilus.		
REFERENCE	1 (bases 1 to 13818)		
AUTHORS	Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A.,		
	Kirkness,E.F., Kerlavage,A.R., Bult,C.J., Tomb,J., Dougherty,B.A.,		
	Merrick,J.D., McKenney,K., Sutton,G.G., Fitzhugh,W., Fields,C.A.,		
	Gocayne,J.D., Scott,J.D., Shetty,R., Liu,L.I., Glodek,A.,		
	Kelly,J.M., Weidman,J.F., Phillips,C.A., Spriggs,T., Hedblom,E.,		
	Colton,M.D., Utterback,T., Hanna,M.C., Nguyen,D.T., Saudke,D.M.,		
	Georhagen,R.C., Fine,L.D., Feltsman,J.L., Fuhrman,J.L.,		
	Georhagen,N.S., Gnehm,C.L., McDonald,L.A., Small,K.V., Fraser,C.M.,		
	Smith,H.O. and Venter,J.C.		
TITLE	Whole-genome random sequencing and assembly of Haemophilus		
JOURNAL	Influenzae Rd		
MEDLINE	Science 269 (5223), 496-512 (1995)		
REFERENCE	95350630		
AUTHORS	2 (bases 1 to 13818)		
AUTHORS	Tatusov,R.L., Muehlegan,A.R., Bork,P., Brown,N.P., Hayes,W.S.,		
	Borodovsky,M., Rudd,K.E. and Koonin,E.V.		
TITLE	Metabolism and evolution of Haemophilus influenzae deduced from a		
	whole-genome comparison with Escherichia coli		
JOURNAL	Curr. Biol. 6 (3), 279-291 (1996)		
MEDLINE	96398784		
REFERENCE	3 (bases 1 to 13818)		
AUTHORS	White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-JUN-1995) The Institute for Genomic Research, 9712		
AUTHORS	Medical Center Dr, Rockville, MD 20850, USA		
AUTHORS	4 (bases 1 to 13818)		
TITLE	White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.		
JOURNAL	Direct Submission		
REMARK	Submitted (27-SEP-1997) The Institute for Genomic Research, 9712		
	Medical Center Dr, Rockville, MD 20850, USA		
	The H. influenzae sequence has been updated by R. Fleischmann, New		

improved, and a number of frame shifts have been corrected. gratefully acknowledge the work of Tatsuo et. al. We have incorporated their annotation into the /notes fields of the corresponding H. influenzae genes
5 (bases 1 to 13818)

Db 841 GGGATATTACGGGAAGCAACCCCTTTGAAGCAAGAAATGGTTTCATTTGGTTCCG 900
Qy 913 aaagattggtctatgatgaatcttctgtagagataccattcagttatcttaagaa 972
Db 901 AAATCTTATTT---TGATCAAAATTTTTCGAAGAGATTTACATACATCACTTTACGCCA 957
Qy 973 cgtcaaatgggaatactcctttaaagaataatgacaggaataatcaatgcc 1032
Db 958 GCTGGTAATGGAGTGCATCAATAGTGAATGATTAATGTCAGGGGCTATTAATCTCAG 1017
Qy 1033 aaactgaacaactctcctgctatagattaaacaacagacccgttcaattgitta 1092
Db 1018 AAATCAGGAATTCATTCAGAAATTTAAA-----TTACGTTAGCAAT 1059
Qy 1093 gttcttatacagagacagcaagaaacctgttatactatgcgc---gggtgtcgaac 1149
Db 1060 ATGATCTTACCTTTAAAGAGAGATTAAGTTCATTAATCTAGATATGACGGACTAAT 1119
Qy 1150 agttaagcccaagcigtatataatgaataatcttcttattatgacgaagaagaagc 1209
Db 1120 ATTATTTCTCCAGCTTTAACAATGAGAAACGCTATATTTATGATCAAAAACAGSA 1179
Qy 1210 gaattgatactatcaacaacaatcaatcaagtgctgtagaataatcttcaagagat 1269
Db 1180 TCATTAATCTTCGATCTGACATTAACCAAGGGGGGTGCTTATTTTGAAGGCTAAT 1239
Qy 1270 tttaagctcgcgcgtgaataatgaacttgcgaagcgcggtgttataatcagtgaa 1329
Db 1240 TTTACAGTATCTCCAAATTTTACCAAACTTGGCAAGGAGCTGATCATATGTAATGTA 1299
Qy 1330 gacagtlacgcttacttgaagaatgaacgctgcgcgaacacgacgcctgtccaaatcgc 1389
Db 1300 AATACACCGCTACTTTGGAATGTAATGGCTGGAACATGATGCACTTTCTAAATTTGGT 1359
Qy 1330 aaagcaacgctgcgaactcaagccaaaggggaaacaaagcgtcgatcgctggcgac 1449
Db 1360 AAAGGACATTTGCGACGTTCAAGCCAAAGGGGAAATTAAGGTTCAATCGCTTGAAGCAT 1419
Qy 1450 ggtacagtcacttctgatacagcagcagcagcagcagcagcagcagcagcagcagcagc 1509
Db 1420 GGTAAAGTCATTTTGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1479
Qy 1510 atcgcttgcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1569
Db 1480 ATGGCTTGGTTAGGCGCAGAGGAGCTGTTCAATTAAGATGAATTAACCAATTTGTTAC 1539
Qy 1570 gacaaactatctcgcttgc 1629
Db 1540 GATTAATTTTATTTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1599
Qy 1630 ttcacagctatcaaaaatacagcagcagcagcagcagcagcagcagcagcagcagcagc 1689
Db 1600 TTTAAACGATTCGAAATTAAGGAGCAGGAGGAGCAGTATGTAATTAACCAATTTGTTAC 1659
Qy 1690 gaatcacgcgttaccattacagcagcagcagcagcagcagcagcagcagcagcagcagc 1743
Db 1660 GCGGTATATGTCATATTAATCTGGAACGAAACGATTTGTTCTACATTAATGAATATAT 1719
Qy 1744 aacagcttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1803
Db 1720 AATAAATCTGATTAAGAAAGAAATGCTCAACGAGTTGCTTGGCCAAACAGATTA 1779
Qy 1804 accaagaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1863
Db 1780 AATAAACAATGAGGCGATTAATTAATCTTATTTAACAACACACAGAGATGCTACTTTG 1839
Qy 1864 ctgcttccgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1923
Db 1840 CTACTTTAGGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1899
Qy 1924 ttcagcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1983
Db 1900 TTACAGCGGATGACGACACGACGCTACATCAATTTAATTAATTAATTAATTAATTAAT 1959

Qy 1984 gaagcattccctcgc 2043
Db 1960 GAAGGATATCCCAAGAGGCAATTTGCTGGATACAGATTTGATCAACGCTACATTTAA 2019
Qy 2044 gcggaacattcccaaatgaagaagcagcagcagcagcagcagcagcagcagcagcagcagc 2103
Db 2020 GCTGAATCTTCGAAATTAAGCGGAGAGTGGGTTGCTTCGCAATGTTCTTCAAT 2079
Qy 2104 aaagcagcttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2163
Db 2080 GAGGGAATTTGGACATGACATATATGCAATGCAATTTGCTGTTGTCGAATTC 2139
Qy 2164 agccacaacatctgacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2223
Db 2140 CAAATATCCATTTGACGCGGCTGAGTTCAGATTCAGATTCAGATTCAGATTCAGATTC 2199
Qy 2224 attaccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2283
Db 2200 TTACCGATACAAAGTATTATTTATTCATTCACAAACCAATTCATTCATTCATTCAT 2259
Qy 2284 ctgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2343
Db 2260 TTAATGATTAATGCAACGCGCAATGTTAAAGTTTGAAGAACTTATGCAATGTCAC 2319
Qy 2344 gcaaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2403
Db 2320 TTACCAATTCACGCAATTTACATTAACCAACCAATTCAGCAATTCAGCAATTCAG 2378
Qy 2404 ctgcgtggaatgcccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2463
Db 2379 ----- 2378
Qy 2464 tgggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2523
Db 2379 -----ACTT 2382
Qy 2524 tccgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2583
Db 2383 TCCGACATTAATTCACGCGGATGATGCAATTCGCAATTCGCAATTCGCAATTCGCAAT 2442
Qy 2584 gataagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2643
Db 2443 GATTCAGCTCAATTTCTTTAAATAACAGCCATTTTCCACCAATTAAGGAGCA 2502
Qy 2644 gatacgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2703
Db 2503 GGCACACAGTGAAGTGAATAATGCAATTCGCAATTCGCAATTCGCAATTCGCAATTC 2562
Qy 2704 aattcaaccttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2763
Db 2563 AATTAAACCTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2608
Qy 2764 ggggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2823
Db 2609 -----CAGTACGTCACAAATTAATGACGCGCGCTTCAAT----- 2644
Qy 2824 tccctatttcggttacacgc 2883
Db 2645 ---TAGACACGGAACCAAGCCCAATCATGCGGAGCAATGTTTCAACACATTGACAGTA 2700
Qy 2884 aacgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2943
Db 2701 AATGGAATTTAGTGGGCAAGGACATTCATTTACTTCAATCTTATTTGGCTATTA 2760
Qy 2944 agcgacaaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3003
Db 2761 AGCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2820
Qy 3004 accgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3063
Db 2821 ACAGGAAGAAGACCGCAACCTTGAAGCAATTAATTTGTTGAAGCAAGTAATATCA 2880

[illegible]

Db	3961	TATTAATGCTGGCAATTCGAGTTGATTATTTACATTACTCCGACAGATATATATCAGGGTTAAG	4020
Qy	4153	ccattatltgagccctgtcctaataccgaatlgccgcttcgggcaaatgcgaaacacgcgtcaat	4212
Db	4021	CCTTATTTCTTGCTGCATATATATGTGATGTTCAAAGCGTAAACGTACACCAACACCGTAAT	4080
Qy	4213	accgcgcgatltgctcagaatalttcgcaaaaccgcgcagtlcggaatlyggcggttaaacgcc	4272
Db	4081	CTCACGGTCTTGCAACACCATTTGGACGTTATTTGGCAAAAGACAGTGGGCTTTAAAGGCA	4140
Qy	4273	gaatlcnaaggtltcacgcgtctgcctcccaacgcctgcccgcgcgcaaaagtcgccgaactgaa	4332
Db	4141	GAATTTTTCATTTTCCAAATTTTCCGCTTTATCTCAAAATCTCAAGGTTCACCACTGGCC	4200
Qy	4333	ggcgcaacacagcgcgggacalcaaatagctacccgctgtgtaa	4374
Db	4201	AAACACGCAAAATGTGGCGCGTAATTTGGCTATTCGTTGGTTAA	4242

RESULT 14	
ARI57320	
4319 bp	DNA
	linear
	PAT 17-OCT-2001

VERSION	AK13/520.1	GI:10210202
KEYWORDS		
SOURCE	Unknown.	

ORGANISM	Unknown. unclassified.
REFERENCE	1 (bases 1 to 4319)

AUTHORS St. Geme, J. W. III and Falkow, S.
TITLE Haemophilus adherence and penetration proteins
ADDRESS Department of Microbiology, University of Michigan, 1137 TUN-2001, Ann Arbor, MI 48106-2001.

JOURNAL	Patent:	US 0245537-A	1 12 00
FEATURES			
source			
	Location/Qualifiers		
	1. .4319		

	/organism="unknown"			
BASE COUNT	1497 a	776 c	891 g	1155 t
ORIGIN				

DATE: 06-08-2000 13:40.4. DB: 6. Length: 4319.

Query Match	26.06;	Score	1243.74;	DB	0;	Length	1227;
Best Local Similarity	58.88;	Pred. No.	8e-271;				
Matches 2565; Conservative	0;	Mismatches	1556;	Indels	264;	Gaps	13

Qy 73 cctgcttacttaagcatalgtcgttcgttcggcattcttccccaagcctggtcggtacac 132

DB 81 CTTAATTTTAAACCGCTTGCAITTCATTAGGAGATACATCGCAAGCGTGGCGTGTCAC 140

0y 133 acttatttcgacacaaactacacatctcgcgacatttgcgaaacaaagccaaagcc
 141 acttatTTGGGATGATTACCAATATTATCGTGATTTTCCCGACAATTAAGGAAGTTC 200

Qy 193 gcagtcgaggcgaaagatatctaggtttacacacaaaaaggaggatttgctgcgcaaatca 252

Db 201 ACAGTGGGGCTCAAAATATTAAAGTTTAAACAACAAAGGGCAATTAGTTGGCACAATCA 260

Db 261 ATGACAAAGCCCCGATGATTGATTTTCTAGTGTACGTACGTAACGGCGTGCACGCTTG 320

313 gtggcgatcataatatgttgagcgtggcacataacggcgcgtatatcaacaagttgattt 372

Db 321 GTTGAATTCATATATTTGTAGCGTGGCACATAACGTGATATACAGATGTTGATTTT 380

373 ggtcggaaggaatcccgatcaacatcgcttctacatadadadcttgcgaatggac
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684
 685
 686
 687
 688
 689
 690
 691
 692
 693
 694
 695
 696
 697
 698
 699
 700
 701
 702
 703
 704
 705
 706
 707
 708
 709
 710
 711
 712
 713
 714
 715
 716
 717
 718
 719
 720
 721
 722
 723
 724
 725
 726
 727
 728
 729
 730
 731
 732
 733
 734
 735
 736
 737
 738
 739
 740
 741
 742
 743
 744
 745
 746
 747
 748
 749
 750
 751
 752
 753
 754
 755
 756
 757
 758
 759
 760
 761
 762
 763
 764
 765
 766
 767
 768
 769
 770
 771
 772
 773
 774
 775
 776
 777
 778
 779
 780
 781
 782
 783
 784
 785
 786
 787
 788
 789
 790
 791
 792
 793
 794
 795
 796
 797
 798
 799
 800
 801
 802
 803
 804
 805
 806
 807
 808
 809
 810
 811
 812
 813
 814
 815
 816
 817
 818
 819
 820
 821
 822
 823
 824
 825
 826
 827
 828
 829
 830
 831
 832
 833
 834
 835
 836
 837
 838
 839
 840
 841
 842
 843
 844
 845
 846
 847
 848
 849
 850
 851
 852
 853
 854
 855
 856
 857
 858
 859
 860
 861
 862
 863
 864
 865
 866
 867
 868
 869
 870
 871
 872
 873
 874
 875
 876
 877
 878
 879
 880
 881
 882
 883
 884
 885
 886
 887
 888
 889
 890
 891
 892
 893
 894
 895
 896
 897
 898
 899
 900
 901
 902
 903
 904
 905
 906
 907
 908

433 aattataaagcaggactaaagccatccctatgtgcgcgattatcatatgcccgtttg 492

Db 441 AACTACAAAAG--ATAATTACATCCTTATGAGAGCATTTACCAATATCCACGATTA 497

CDS

```

<1..>576
/codon_start=1
/transl_table=11
/product="Iga protease"
/protein_id="CA04965.1"
/db_xref="GI:3152400"
/db_xref="SPTREMBL:O69746"
/translation="GKFAVGAKDIEVYNNKGELVGSMTKAPMIDFVSVRNGVALY
GDOYIVSAHNGVNNVDGAGSNPDHREFOYIKRNNYKPDNSHPYNGDYHMRRL
HKFVTDAPVEMTGDMRGNTYSDEKYEPRVRIGSGHHYWRIDDDKHGDLSTSGAWLI
GNTHMGGMGNNGVYSLSGDVRHANDYGMPLI"
repeat_region
88..95
/rpl_type=INVERTED
repeat_region
139..148
BASE COUNT
163 a 125 c 156 g 134 t
ORIGIN

```

Query Match 7.3%; Score 318.8; DB 1; Length 578;
Best Local Similarity 83.3%; Pred. No. 4.1e-61;
Matches 375; Conservative 0; Mismatches 72; Indels 3; Gaps 1;

```

QY 184 ggcgaattgcagtcgagcgaagaattgagggtttacacaaagaaggaggttgc 243
Db 1 GCGAAGTTGCGAGTCGGGCGCAAGATTTGAGGTTTACACAAAAAGGGAGTTGGTC 60
QY 244 ggcgaatcaatgacaaaagcccgatgattgatttctgtgtgtcgcgtaacgagcgtg 303
Db 61 GCGAATCAATGACAAAAAGCCCGCATGATTTTTCGCTGTCGCGTAAAGCGCGTG 120
QY 304 ggcgcattggtggcgatcaatatgttgagcgtggtgcataaaggcggtataaac 363
Db 121 GCGCATTTGGGGGCGATATATTTGTGACGCTGACATACGGCGCTATTAACAAT 180
QY 364 gtgattttgttgcggaagaagaatcccgatcaacatcgctttactataaatttgcg 423
Db 181 GTGATTTTGTGCGAGAGGGAAGCAATCCGATCAGACCGCTTTCTTACCAATTTGTG 240
QY 424 aaagcgaataattaagaagcgaagcgaagcgaatcccttaagcgcgtatcatatg 483
Db 241 AAAAGAATCACTACAAACCTGACATTTCA--CACCTTACACGGCGATTAACATATG 297
QY 484 ccgcgttcgataattgtcacagatgcagaacctgttgaatgacagttatatg 543
Db 298 CCGGCGCTGCATAGTTCGTAAACGATCAGAACTGTGGAATGACGGGTACATGAGG 357
QY 544 ggcgcgaatatatcgatlcataaataatcacctgacggtgttcgtatggtgcagcgag 603
Db 358 GGGAAATACCTATTCGATAAAGAAAAATATCCGAGCGTGTCCGATGGGCTCAGAGAC 417
QY 604 caatatggcagctcgtatgaatgagccc 633
Db 418 CACTATTGGCGTTATGATGACAAACAC 447

```

RESULT 16
NLIGA3354 584 bp DNA linear BCT 21-MAY-1998
LOCUS NLIGA3354
DEFINITION *Neisseria lactamica* Iga protease homologue gene, isolate NL3354.
ACCESSION AJ001739
VERSION AJ001739.1 GI:3152397
KEYWORDS Iga1 protease.
SOURCE *Neisseria lactamica*.
ORGANISM *Neisseria lactamica*.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 584)
AUTHORS Sayers, J.R.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1997) Sayers J.R., Division of Molecular
Medicine, University of Sheffield, Royal Hallamshire Hospital,
Sheffield, S. Yorks, S10 2JF, ENGLAND
REFERENCE 2 (bases 1 to 584)

AUTHORS Sayers, J.R., Vitowski, S. and Read, R.
TITLE Strains of *Neisseria lactamica* carry an Iga1 protease gene
JOURNAL homologue and some express Iga1 protease activity
FEATURES
source Location/Qualifiers
1..584
/organism="Neisseria lactamica"
/lsolate="NL3354"
/db_xref="taxon:486"
<1..>582
/codon_start=1
/transl_table=11
/product="Iga1 protease"
/protein_id="CA04964.1"
/db_xref="GI:3152398"
/db_xref="SPTREMBL:O69747"
/translation="FAVGAKDIEVYNNKGELVGSMTKAPMIDFVSVRNGVALYD
OYIVSAHNGVNNVDGAGSNPDHREFOYIKRNNYKPDNSHPYNGDYHMRRLK
FVTDAPVEMTGDMRGNTYSDEKYEPRVRIGSGHHYWRIDDDKHGDLSTSGAWLI
GNTHMGGMGNNGVYSLSGDVRHANDYGMPLIAG"

CDS

```

repeat_region
82..88
/rpl_type=INVERTED
repeat_region
134..142
BASE COUNT
163 a 127 c 159 g 135 t
ORIGIN

```

Query Match 7.2%; Score 313.6; DB 1; Length 584;
Best Local Similarity 83.6%; Pred. No. 6e-60;
Matches 368; Conservative 0; Mismatches 69; Indels 3; Gaps 1;

```

QY 190 ttgcagtcgggcgaagaatattgagttacacaaagaaggaggttgcgcaaa 249
Db 1 TTGCACTCGCGGGGGAAGATTTGAGGTTTACACAAAAAGGAGTTGTGGCAAA 60
QY 250 tcaatgacaaaagcccgatgatttctgtgtgtcgtgtaacggtgcgcgca 309
Db 61 TCAATGACAAAAGCCCGCATGATTTTTCGCTGTCGCGTAAACGGCGTGGCGCA 120
QY 310 ttgttggtgcgataatatgttgagcgttgacataacgagcgtataaacgttgc 369
Db 121 TTGCTGGCGATCAATATTTGTGAGCGTGACATTAACGGCGCTATTAACAATGTGAT 180
QY 370 ttgtgtgcggaagaagaatcccgatcaacatcgctttactataaattgtgaacgg 429
Db 181 TTGTGTCAGAGGGAAGCAATCCGATCAGACCGCTTTCTTACCAATTTGTGAAAAG 240
QY 430 aataatataagcaggttaagcgaatcccttaatgagcgtatcatatgcccgt 489
Db 241 AATAACTACAAACCTGACATTTCA--CACCTTACAAACGGCGATTAACATATGCCGCGC 297
QY 490 ttgcataattgtcacagatgcagaacctgttgaatgacagttatatgagtgaggcg 549
Db 298 CTCGATTAAGTTCGTAAACGATCAGAACTGTGGAATGACGGGTGACATGAGGGGAT 357
QY 550 aaatatatcgatcaaaataatcacctgacggtgttcgtatggtgcagcgacaatat 609
Db 358 ACCTATTCCGATTAAGAAAAATATCCGAGCGTGTCCGATGGGCTCAGGACACCACTAT 417
QY 610 tggcagctcgtatgaatga 629
Db 418 TGCGTTATGATGATGACGA 437

```

RESULT 17
NLIGA823 572 bp DNA linear BCT 21-MAY-1998
LOCUS NLIGA823
DEFINITION *Neisseria lactamica* Iga protease homologue gene, isolate NL823.
ACCESSION AJ001737
VERSION AJ001737.1 GI:3152389
KEYWORDS Iga1 protease.
SOURCE *Neisseria lactamica*.
ORGANISM *Neisseria lactamica*.

REFERENCE	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
AUTHORS	Neisseria.
TITLE	1 (bases 1 to 572)
JOURNAL	Sayers,J.R. Direct Submission Submitted (20-OCT-1997) Sayers J.R., Division of Molecular Medicine, University of Sheffield, Royal Hallamshire Hospital, Sheffield, S. Yorks, S10 2JF, ENGLAND 2 (bases 1 to 572) Sayers,J.R., Witowski,S. and Read,R. Strains of Neisseria lactamica carry an IgM1 protease gene homologue and some express IgM1 protease activity Unpublished
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
FEATURES	
source	Location/Qualifiers 1..572 /organism="Neisseria lactamica" /isolate="NL823" /db_xref="taxon:486" CDS <1..304 /codon_start=2 /transl_table=1 /product="Iga protease homologue" /protein_id="CAA04962.1" /db_xref="GI:3152390" /db_xref="SPRMBL:069748" /translation="KDIENVYKKCELDICKSMFKAPMDIFSVSRNGVALVGDQIVYSVAHNGGYNNDPCGAEGGNPDQHRPSTQIVKKNNTKPDSNRPNGDYHMPRLPGISS"
repeat_region	44..51 /rpt_type=INVERTED
repeat_region	68..75 /rpt_type=INVERTED
BASE COUNT	158 a 124 c 157 g 133 t
ORIGIN	
Query Match	6.4% Score 280.6; DB 1; Length 572; Best Local Similarity 81.1%; Pred. No.1,7e-52; Matches 353; Conservative 0; Mismatches 74; Indels 8; Gaps 2;
Dn	204 gaagataatgagggttcaacaaaaaaggagtgctgcggcaaatcaatgacaagaac 263
Dn	1 GAAGATATTGAGGTTCACAAACAAAAGGGAGTTGGTGGCAATCATGACAAAAGC 60
Oy	264 ccggatgatgattttcttgtgtgtcgttaagcgctggcgcatgtgtggcgatca 323
Dn	61 CCCGATGATGATTTTCCGTGCTGCCGTACCGCGCGCATGCTGGCGATCA 120
Oy	324 atatatgtgacgtggtcacataaacggtccgtatatacaaacgttgaatttggtgcggaag 383
Dn	121 ATATATTGTGAGCGGGCGCATTAACGGCGGCTATTAACAATGTTGATTGGTGCGGAGGG 180
Oy	384 aagaatcccgatcaaatcgtttacttatataaaatttgaaacgaataatataaagc 443
Dn	181 AGGCATATCCGATGCGACACCCTTTCTTCCAAATTTGTAATAAAGAANAATAACAAACC 240
Oy	444 agggactaaaggccaaccttatlgcggcgatatalatagccgcg-----tttcataaa 498
Dn	241 TGACAAATTC---CACCTTACAAAGGGGAGATTACCATATGCGCGCCGCTGGCATANG 297
Oy	499 ttgtgcacaagtggagaacgtgttaaatgaccagttatgtatgagggcggaataatatic 558
Dn	298 TTCGTAACCGATGCGAACACTCTCCAAATGACGGGTGACATGAGGGGGAATACCTATTTC 357
Oy	559 gatcaanaatattaccctgcagtglttgcgtatttggggaaggaagcaatattggcgatct 618
Dn	358 GATAAAGAAAAATATCCGACGCTGCCGATCCGCTCAGACACACACTAATTTGGGTTAT 417
Oy	619 gatgaagatggccc 633
Dn	418 GATGATGACAAACAC 432

LOCUS	NLGA3293	572 bp	DNA	linear	BCT 21-MAY-1998
DEFINITION	Nisseria lactamica IgA protease homologue gene, isolate NL3293.				
ACCESSION	AU001178				
VERSION	AU001738.1	GI:3152385			
KEYWORDS	IgA1 protease.				
SOURCE	Nisseria lactamica.				
ORGANISM	Nisseria lactamica				
	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;				
	Neisseria.				
REFERENCE	1 (bases 1 to 572)				
AUTHORS	Sayers,J.R.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-OCT-1997) Sayers J.R., Division of Molecular Medicine, University of Sheffield, Royal Hallamshire Hospital, Sheffield, S. Yorks, S10 2JF, ENGLAND				
	2 (bases 1 to 572)				
REFERENCE	Sayers,J.R., Witowski,S. and Read,R.				
AUTHORS	Strains of Nisseria lactamica carry an IgA1 protease gene homologue and some express IgA1 protease activity				
TITLE	Unpublished				
JOURNAL					
FEATURES					
source	location/Qualifiers				
	1..572				
	/organism="Neisseria lactamica"				
	/isolate="NL3293"				
	/db_xref="taxon:486"				
	<1..>570				
CDS	/codon_start=1				
	/transl_table=11				
	/product="IgA1 protease"				
	/protein_id="CA04963.1"				
	/db_xref="GI:3152386"				
	/db_xref="SPTRMBL:O69745"				
	/translation="KDIENVKKKEIYKSMFKAPMIDESVSRNGVALVDQYISVAHVGRYDVPFGAGSNPDHREYSYLAKNNKNKDETHPEXDYNPLRKVTFRAPAIPMTSDMDNKITDRKTIERVRISGWFPMNDODKGDVAGAVHYLTAGTTHMOGGAGGSWSLSGDVRHGANGPIPIRAGSS"				
	67..74				
	/rpe_type=INVERTED				
repeat_region	repeat_region	118..127			
BASE COUNT	169 a 119 c 154 g 130 t				
ORIGIN					
	Query Match	6.4%; Score 280.2; DB 1; Length 572;			
	Best Local Similarity	77.7%; Pred. No. 2.le-52;			
	Matches	352; Conservative 0; Mismatches 98; Indels 3; Gaps 1;			
OY	205 aaagatattgaagtcttacacaaacaaaaggaggggttgcgcgcgaatacatgatcaaaagcc	264			
Db	1 AAAGATATTGACGTTTACACAAAAAAGSGGAGTTGGTCGGAAATCATGCATAAACCC	60			
OY	265 ccgatgatgatcttttcctcgtggtgcgcgttaacggcgctggcgcatcttggtggagtaa	324			
Db	61 CCGATGATGATTTTCTGTGGTATCAGCGTAACGCCGTGGCGCATTTGGTGCGATCAA	120			
OY	325 tatattcgtgagcgtgcacataacgacgcgcataacaaacgttgtattcttggtgcgaaaga	384			
Db	121 TATATTGTGAGCGTGGCGATATCTAGGTTATCCGATGTTGTTGGTGGAGGGA	180			
OY	385 agaaatcccgatcaacalcgtttactactataaaattcgtgaacggaaataattataagca	444			
Db	181 ACCAATCCCAGTACCACACCGTTTTTCTTATAAGATTGAAAACGGAAATAATTATTAAC	240			
OY	445 ggagctaagagcacaccatagcgcgcgacataccaatalcgcgcgcttgcatataattctc	504			
Db	241 G---ATGAAGAAGCATCTTATATGAGAAAGATTACCAACAACCCACGCTTACATPAAATTGT	297			
OY	505 acagatgcagaacctggttgaaatgcccgatatcatgatgtggcgcgaaatatatcatcaa	564			
Db	298 ACGGAAGCCGCCCATGATGATGACTTCTGATATGAGACGGCAACAATATPACAGATAGC	357			
OY	565 ataattaacctgcacgctgctcgttatcttggtggagcgcgacatatctgcgtctgataa	624			

OY	1675	cacatcataagcaaaagaatccaccggttacatttaagggaataagatatggctaacacc	1734
D8	1405	CACACAACAGCAGCGCCGCTCCACATCACATTAAACGGCAATAACTTAATTACCCTTCCT	1464
OY	1735	gcacaat	1740
D8	1465	CANAAT	1470
RESULT	25		
NMIGAG117			
LOCUS		1525 bp	DNA
DEFINITION		N.meningitidis lga gene (NG117).	Linear BCT 14-MAR-1995
ACCESSION	X82483		
VERSION	X82483.1	GI:732881	
KEYWORDS	lga gene; lgal protease.		
SOURCE	Neisseria meningitidis.		
ORGANISM	Neisseria meningitidis		
REFERENCE	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;		
AUTHORS	Neisseria.		
TITLE	1 (bases 1 to 1525)		
JOURNAL	Lomholt H., Poulsen K. and Kilian M.		
MEDLINE	Comparative characterization of the lga gene encoding lgal protease		
REFERENCE	In Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus		
AUTHORS	Influenzae		
TITLE	Mol. Microbiol. 15 (3), 495-506 (1995)		
JOURNAL	95302961		
MEDLINE	2 (bases 1 to 1525)		
REFERENCE	Lomholt H.		
AUTHORS	Direct Submission		
TITLE	Submitted (19-OCT-1994) H. Lomholt, Dept. of Medical Microbiology,		
JOURNAL	The Barklin Building, University of Aarhus, 8000 C Aarhus C,		
FEATURES	DENMARK		
SOURCE	location/Qualifiers		
gene	1..1525		
CDS	/organism="Neisseria meningitidis"		
	/strain="NG117"		
	/db_xref="taxon:487"		
	1..1525		
	/gene="lga"		
	<1..>.1525		
	/gene="lga"		
	/codon_start=1		
	/transl_table=1		
	/product="lgal protease"		
	/protein_id="CAA57866.1"		
	/db_xref="GI:732882"		
	/db_xref="SPTRMBL:O511171"		
	/translation="ATDLSVKKKGOGINTGNALSNVPMIDFSVDVNKRITAVPQFPA VSVKRAKAENVHFFYGQYNHNNDVADKENERYVEONNRPBHKMGASNLGRLEDYNNR ARENEFEVAPAIAPTDAAGGGLDTTKDKNRSSPRVRAGAQOLYERKAYHOEENGEG KYDLRLSOAARYAIAGTGPYKDINDIQTNTBEGLIFGFNHNIHSABELKOALSODALL NVGVLDGSSEPLFAFDKORNMWVELGTVDYANGVGKSKSQEWNIYKEFAFKIKORD AGTIGNREHHWNITPEFGNSRKSGSTAVVLANNERDANNQNVTEDNGTLVDNONIC GAGGEFPFGDYTVKGANDDIWLGAIGIVADAKKVWMGVKKPNGBRLRIKGLTEIRI GVWGGOGLKGYDDGVILINCOADADAKKVOAFESOVGITVSRGLVLTNSSNQINPDMLYLE GFRGRIDLANSMDLTFEHHRNRYNBDEGARIVRNHTDRASTITLGKSLLTPAAPONSLEYE		
BASE COUNT	507 a	318 c	334 g
ORIGIN			366 t
Query Match	4.5%	Score 195.6:	DB 1: Length 1525;
Best Local Similarity	61.2%:	Pred: No. 2, 9e+33:	
Matches 334:	Conservative 0;	Mismatches 209;	Indels 3; Gaps 1
OY	1198	gaaggaanaagcgaattgatctacttaccagaacaatcaatcaatgaagtgctcgaggaattat	1257
D8	961	GAAAGACAAGCGCACTTTGGATTGGATCATCAAAACATCATCAACCAAGCGGCGGCTCTGTTT	1020
OY	1258	tccaagagagattttaagtcctcgcctcgaaataaaga---acttggcaaagcgcgggcg	1314
D8	1021	TTCCAAGAAGCGATTACAGATGCACAAAGTGCAAAATPATATACATCACTTGCTAGGTTGGG	1080

QY	1315	gtctatatacagtagaagacagtagtaccggttaacttcttggaagttaaagcgctgycgcaaacgaccgc	1374
Db	1081	ATTGATGTGTGGCGACGGCCAAAAGTGGTTTGGCAAGTCAAAATVCCGAATGGCGACAGA	1140
QY	1375	ctgcgccaaatccgacaagcgacgctgcagcttcaagccaaaggagaaaccgaagctcg	1434
Db	1141	TTGGCAAAAATTCGGCCAAAGGCACATTAGAAATTAAGCGCACAGGCGTTAAACCAAGGGCMA	1200
QY	1435	atcagcgtgycgacgagctacgtaacttlttgataagcagcgacagcagataaaggcaaaaa	1494
Db	1201	TTAAAAGTCGGCGACGGGTATCTTGATGATCAACCAACCGATGCGGACAAAAAAGTC	1260
QY	1445	caagccttaagtgaatgcgttctgttcaagcgagcggtacgycgtcaactgaatgacgat	1554
Db	1261	CAGGCTTCTCCCAAGTGGGCACTTGTCTACGCGACGGACGGGTACATTGTGTAAATAGTTC	1320
QY	1555	aatcagttcaaccccgacaaactctatttgcgcttcgycgacgacgcttggatlltaaac	1614
Db	1321	AATCACTTAAATCCCGATAACTTATATTTCGTTCCGTGCGGTCGTTTGGATGCCAT	1380
QY	1615	gggacatcgcttctgttccacgtaattcaaaataacccgaatgaagggcgatgattgcaac	1674
Db	1381	GGCAATGACTTGACTTTTGAACACACTCCGCCAACGTGAGTAGAAGGGCGGATGTCAAC	1440
QY	1675	cacaaatcaagaacaaagaatccacgcttaccattacaggaataaagatactgtctacaac	1734
Db	1441	CACAAACAGACCGCGCTCCACACATTCACATTAAACGGCGAAATCAATTATTAACCGCTCCT	1500
QY	1735	ggcaat 1740	
Db	1501	CAAAAT 1506	

```

RESULT 26
NGIGAPRO          NGIGAPRO          4899 bp      DNA      linear      BCT 21-MAR-1995
LOCUS              Neisseria gonorrhoeae gene for Iga protease.
DEFINITION          X04835
ACCESSION            X04835
VERSION              X04835.1 GI:44868
KEYWORDS             Iga protease; Inverted repeat; protease; secreted protein; signal
                     peptide.
SOURCE               Neisseria gonorrhoeae.
ORGANISM             Neisseria gonorrhoeae
                     Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
                     Neisseria.
REFERENCE            1 (bases 1 to 4899)
AUTHORS              Pohlnert,J., Halter,R., Beyreuther,K. and Meyer,T.F.
TITLE                Gene structure and extracellular secretion of Neisseria gonorrhoeae
                     Iga protease
JOURNAL              Nature 325 (6103), 458-462 (1987)
MEDLINE              87115823
FEATURES
source              Location/Qualifiers
                     1..4899
                     /organism="Neisseria gonorrhoeae"
                     /strain="M11"
                     /db_xref="taxon:485"
                     /clone="p1p100"
                     34..39
                     /note="put. -43 region"
                     43..48
                     /note="put. -35 region"
                     63..68
                     /note="put. -10 region"
                     87..97
                     /note="put. rRNA binding region"
                     104..4702
                     /note="Iga protease precursor; 169K (AA -27 to 1505)"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAA28538.1"
                     /db_xref="GI:44869"
                     /db_xref="SWISS-PROT:P09790"

```

```

/translation="MKAKRKINAIISLIFLAVALNPYSEALVRDQVDFRFAE
NKKFPEVATDLSVKKRKGONIGNALSNVPMIDFVSADVNRKRIATVVDPOYAVSKHA
KAEVHEFYGOYNGHNDVADKRENEYRVONNPEPRKAWGKSLGLEDYENMRKPFKE
VTEVAPIAPTDAGGIDTCKDKNRESEFPICAGRPOLYKGYHOGEGNGYDLRDL
SOAYRYAIACTPKDINTDQTMTEBELLFGNNKQYSAEELKQALSODALTYNYGUG
DSGSLPFAFDKONMOWFLGTIDYMAGYKSKSMOEWNIYKKEFADIKONDNAGTYGK
NEGHMKTTGTNSHIGSTAVRLANNENGDANGONVFEENGTVLVONINOGAGGLFE
KGDYVKGANNIDITWLGAGIDVADKRYVOWNPNGDRIAKIGKLTLEINGVONG
OLVGDGTVLINOKADADKVOAFSGVIGSGGTIVLSSNOINPDNLYTFGGVNG
DANGNDLTFEIRNDEGARIVNHTDASTLTGTSLITNNSVHSIONDYED
DYSYRRPRPPIPOGKDLKYKRYALSGGLNAMPENGAENNDIMPEWTOBE
ARKNMMHNKRNRIIDGEGFDEENGKNGALNINFNKSAQRELLTGALNKKI
SVTGGVLLSGRPHARDPVFKSSRKRAHNSKNEVFEEDMINKRTKAATINKI
SAFSSGRNVSDITANTITATDANKVNLGTRNGDEVCVRSIDYGTVCNTGNLSKALN
SFDATRINGNVNLNOMNALVILGKAALMGKIQOGGNSRVSINOSKMLTGDSQVHNS
LADSHILNNASDAOSANKYHTIKINHLGNGHFHYLTDLAKMLGDKVLVKSASGHY
OLAVONKTEGPENEGIDLPDASSVODRSRLFEVLAHNYVDLARTIKTENGITRLY
NRYAGNGRPKPAPSPANTASOAKATOTDGOIAKPONIYVAPSPONAKOEAELR
OGAKAEOVKROQAEEKVAROKDEEAKRAFIAROGEARAKAELAKOKAEARK
ABELAROKAEEASHONAKPKRRRRRIILPRPAPVPSIDTDYAKNSESSTNLARV
ITRMGRELINDYEIEPLEDEBAEEROKATOFHKSRRKRAIISSEPSDEDSASY
STSDKHPDONTLEHEKVEVETAGIQPRAAGPRTQAAQADAVSTINTNSALSDAMASTSI
LIDTGAYLTRHIAOKSRADAENKSNVMSNTGYRADYASAOYRRFSKRTOIGDIRS
LSENMIOIGVLTYSOHTFDGAGKNTFVOANLYGKYYLNDAMYADIGAOSVLSRS
LOTOOKANFRNTSIOTGLTIGNTLKIINOFEIYASGIRYSLSSADYIKLGDGSVYKS
MAVTLTLAGIDFAVRKVGMLVYKPLLSAAYFANYGKGVNWGKSFAYKADNOQOTS
AGVALLYRVNLTNVNSITKQOLEKOKSGOIKIOIRP"
104..184
sig_peptide
185..4699
mat_peptide
/product="mature Iga protease (AA 1-1505)"
misc_feature
3061..3062
/note="autoproteolytic site for Iga protease"
misc_feature
3157..3158
/note="autoproteolytic site for Iga protease"
misc_feature
3466..3467
/note="autoproteolytic site for Iga protease"
repeat_unit
4725..4739
/note="autopeptidic site for Iga protease"
terminator
4725..4762
/note="inverted repeat A"
repeat_unit
4747..4762
/note="put. transcription termination signal"
repeat_unit
4747..4762
/note="inverted repeat A"
BASE COUNT 1617 a 1104 c 1080 g 1098 t
ORIGIN

```

```

Query Match 4.5%; Score 195.6; DB 1; Length 4899;
Best Local Similarity 62.3%; Pred. No. 3.5e-33;
Matches 344; Conservative 0; Mismatches 199; Indels 9; Gaps 2;

```

```

QY 1167 gaataatgggaataattcccttattgaaggaagaaggaattgatactaccag 1226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1192 GAACAACGGCAAAATGTTACCTTT-----GAAGACACAGCGCACTTGGTATTGAATCA 1245

QY 1227 caacatcaatcaaggtgctgagattatattcccaagagattttacggtctgcctga 1286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1246 AACATCAACCAAGCGCGCGCGCTGTTTTCMAAGCGGATTACACAGTCAAAAGCGC 1305

QY 1287 aaataacga--acttggcaagcgcggtctcatatcaatgaagacagttac 1343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1306 AATTAATGACATCACTTGGTTAGGTGGGAGATTGATGTCGACGCAAAAAGTCTGT 1365

QY 1344 ttggaaatgaagcggtgcaaacgacgctgtccaaataatggcaaacgaagcgctga 1403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1366 TTGGCAAGTCAAAATCCGATGCGACAGATTGGCAAAAATGCGCAAGGCACTTTTGA 1425

QY 1404 cgttcaagcaaaaggaagcgtcatcagcgttgagcgagcgttcaagttac 1463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1426 AATTAACGGGACAGCGGCTTAACCAAGGCAATTAAGGTGGGACGCTTACGTTATTCT 1485

QY 1464 ggttcgacgagcagcagataaagcaaacagccttaagaaatcggtctggtcag 1523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1486 GAATCAAAAAGCGGATGCGCAAAAAGTTCAGGCTTCTCCCAAGTGGCATTTGTCAG 1545

```

```

QY 1524 cggcagggatcggtgcaactgaatgcgaataatcagttcaaccgccgacaacatttt 1583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1546 CGGACGGCGGTACATTTGATTAATAGTCAATCAGATTATGCCGATTAACCTATNTT 1605

QY 1584 cggcttcggcgggcgagcgtttgatttaaacgggcatcgcttcgttccacgattca 1643
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1606 CGGTTCCGTCGGCGGTGCTTGGATGCGCAAGGCAATGACTTGACTTTGAACACATCCG 1665

QY 1644 aaataccatgaagggcgatgattgtcgaaccacatcaaacagaagaatccacgttac 1703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1666 CAACGTGATGATGAAGCGCGCGCATTTGACACGACACAGACAGCAGCGCTCCACATTCAC 1725

QY 1704 cattaagagca 1715
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1726 ACTAACGGGTAA 1737

RESULT 27
LOCUS A02796 4899 bp DNA linear PAT 10-MAY-1993
DEFINITION N.gonorrhoeae gene for iga-protease precursor.
ACCESSION A02796 GI:345152
VERSION 1ga protease.
KEYWORDS Neisseria gonorrhoeae.
SOURCE Neisseria gonorrhoeae.
ORGANISM Neisseria gonorrhoeae;
          Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
          Neisseria.
REFERENCE 1 (bases 1 to 4899)
AUTHORS Meyer,T.F., Halter,R. and Pohlner,U.
TITLE Process for recovering proteins by means of gene technology using
          gram-negative host cells
JOURNAL Patent: EP 0254090-A 1 27-JAN-1988;
          Max-Planck-Gesellschaft zur Foerderung der Wissenschaften
FEATURES
    source
        location/Qualifiers
            1..4899
                /organism="Neisseria gonorrhoeae"
                /strain="MS11"
                /db_xref="taxon:485"
                104..4702
                    /codon_start=1
                    /transl_table=11
                    /product="Iga protease"
                    /protein_id="CA00270.1"
                    /db_xref="GI:345153"
                /translation="MKAKRKINAIISLIFLAVALNPYSEALVRDQVDFRFAE
                    NKKFPEVATDLSVKKRKGONIGNALSNVPMIDFVSADVNRKRIATVVDPOYAVSKHA
                    KAEVHEFYGOYNGHNDVADKRENEYRVONNPEPRKAWGKSLGLEDYENMRKPFKE
                    VTEVAPIAPTDAGGIDTCKDKNRESEFPICAGRPOLYKGYHOGEGNGYDLRDL
                    SOAYRYAIACTPKDINTDQTMTEBELLFGNNKQYSAEELKQALSODALTYNYGUG
                    DSGSLPFAFDKONMOWFLGTIDYMAGYKSKSMOEWNIYKKEFADIKONDNAGTYGK
                    NEGHMKTTGTNSHIGSTAVRLANNENGDANGONVFEENGTVLVONINOGAGGLFE
                    KGDYVKGANNIDITWLGAGIDVADKRYVOWNPNGDRIAKIGKLTLEINGVONG
                    OLVDGTGTVLINOKADADKVOAFSGVIGSGGTIVLSSNOINPDNLYTFGGVNG
                    DANGNDLTFEIRNDEGARIVNHTDASTLTGTSLITNNSVHSIONDYED
                    DYSYRRPRPPIPOGKDLKYKRYALSGGLNAMPENGAENNDIMPEWTOBE
                    ARKNMMHNKRNRIIDGEGFDEENGKNGALNINFNKSAQRELLTGALNKKI
                    SVTGGVLLSGRPHARDPVFKSSRKRAHNSKNEVFEEDMINKRTKAATINKI
                    SAFSSGRNVSDITANTITATDANKVNLGTRNGDEVCVRSIDYGTVCNTGNLSKALN
                    SFDATRINGNVNLNOMNALVILGKAALMGKIQOGGNSRVSINOSKMLTGDSQVHNS
                    LADSHILNNASDAOSANKYHTIKINHLGNGHFHYLTDLAKMLGDKVLVKSASGHY
                    OLAVONKTEGPENEGIDLPDASSVODRSRLFEVLAHNYVDLARTIKTENGITRLY
                    NRYAGNGRPKPAPSPANTASOAKATOTDGOIAKPONIYVAPSPONAKOEAELR
                    OGAKAEOVKROQAEEKVAROKDEEAKRAFIAROGEARAKAELAKOKAEARK
                    ABELAROKAEEASHONAKPKRRRRRIILPRPAPVPSIDTDYAKNSESSTNLARV
                    ITRMGRELINDYEIEPLEDEBAEEROKATOFHKSRRKRAIISSEPSDEDSASY
                    STSDKHPDONTLEHEKVEVETAGIQPRAAGPRTQAAQADAVSTINTNSALSDAMASTSI
                    LIDTGAYLTRHIAOKSRADAENKSNVMSNTGYRADYASAOYRRFSKRTOIGDIRS
                    LSENMIOIGVLTYSOHTFDGAGKNTFVOANLYGKYYLNDAMYADIGAOSVLSRS
                    LOTOOKANFRNTSIOTGLTIGNTLKIINOFEIYASGIRYSLSSADYIKLGDGSVYKS
                    MAVTLTLAGIDFAVRKVGMLVYKPLLSAAYFANYGKGVNWGKSFAYKADNOQOTS
                    AGVALLYRVNLTNVNSITKQOLEKOKSGOIKIOIRP"

```

BASE COUNT 1617 a 1105 c 1079 g 1098 t
ORIGIN

Query Match 4.5%; Score 195.6; DB 6; Length 4899;
Best Local Similarity 62.3%; Pred. No. 3.5e-33;
Matches 344; Conservative 0; Mismatches 199; Indels 9; Gaps 2;

1167 gaataatggaagaataatcttcttatttgcagaaggaagaaggaatgacttaccag 1226
1192 GAACAACGGACAAATGTTACCTT-----GAAGACACGGCAGCTTGGATTGATCA 1245
1227 caacatcaatcaagatgctgaggaattatattccaaagaggaatttcggtcgcctga 1286
1246 AATCATCAACCAAGGCGCGGCGCTGTTTTCAAAGGCGATTACACAGTCAAAAGCGC 1305
1287 aaataacgaa----acttgcagaagcgcgcttccatatacagtaagaacagtaaccgttac 1343
1306 AATATATGACATCATTGTTAGTGTGCGGGGATTGATGTCGCCAGCAAAAAGTCT 1365
1344 ttggaagaataaagcgcttgcagaacgacgcttgcagaatcgcgaaggaagcgctgca 1403
1366 TTGGCAAGTCAAAATCCGAATGCGCACAGATTGCGCAAAATCGCAAGGCACTTGTGA 1425
1404 cgttcaagccaaggggaaacccaagggctcgaacgcttgcagcgctgagcggtacagttat 1463
1426 AATTAACGGCACAGGCGGTTAACCAAGGCAATTAAGGTCGCGGATGCGTATGTTCT 1485
1464 ggaatcagcgagcagcagataaagcaaaaacgaagccttgcagtaacgcttgcag 1523
1486 GAATCAAAAAGCGGATGCCAGCAAAAAGTTCAAGCTTTCGCCAAGTCCGCGCATTTGTCAG 1545
1524 cggcaggggagctgctgaactgaatgcggaatacagttcaaccgcgaacactatlt 1583
1546 CGGACGGGATGATGATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1605
1584 cggcttcgagggaggaattggaattaaagggcagcttcgcttcgctccaggtattca 1643
1606 CGGTTTCGGGGGGGCTGTTGGATGCGCAATGGAATGACTTGTGACTTTTGAACACATCCG 1665
1644 aatatacgaagaagggagatgatttgcacaacacatacaagaagaagaatccacgcttca 1703
1666 CAACGTGATGAAGGCGCGGCGCATTTGTCACACACAGACGCGCTCCACATCTAC 1725
1704 cattaacaggcaa 1715
1726 ACTAACGGGTAA 1737

RESULT 28

LOCUS A12416 4899 bp DNA linear PAT 05-OCT-1994
DEFINITION Iga-Protease gene precursor.
ACCESSION A12416
VERSION A12416.1 GI:640585
KEYWORDS
SOURCE Neisseria gonorrhoeae.
ORGANISM Neisseria gonorrhoeae; beta subdivision; Neisseriaceae;
Neisseria.

REFERENCE 1 (bases 1 to 4899)

TITLES

PROCESS FOR THE GENETIC PRODUCTION OF PROTEINS BY USING OF
GRAM-NEGATIVE HOST CELLS
Patent: WO 8800238-A 12 JAN-1988;

JOURNAL

FEATURES

Location/Qualifiers

1..4899

source

/organism="Neisseria gonorrhoeae"

gene

/db_xref="taxon:485"

CDS

/gene="Iga protease precursor"

/gene="Iga protease precursor"

/codon_start=1
/transl_table=11
/protein_id="CAA01030.1"
/db_xref="GI:640586"
/translation="MKARFKINAISLFIPLAATPYSEALVDDVDVDTYERDAE
NKGFEVATDLVYKRGONIGNALSNVPMIDFVADVNKRIATVYDQYAVSKHA
KAEVTEFYGOYNGHNDVADKENEYRVYVONNYEHPKAWGASNLGRLEYNMAREKF
VTEAATPIAPDAGGLDITYKDKNRFSEYRIGAGLOVYKEGVYHQEGNEKGYDLDL
SOAYRYATAGTPYKNDINIDONTNTEGLGEGFNGKSWOMENIYKKEPADIKQHDNAGTVG
DSGSLFAPDQKYNOWPLGTYDVAWAGKSKWOMENIYKKEPADIKQHDNAGTVG
NGEHHMKTGTQNSHIGSTAARLANEGDANNQONTFEDNGTLYLNONINOAGGAF
KGDYTVKANNDITWLAGIDVADKRVYVQKPNBGDLAKITGKTLTNGCVNQG
OLKYDGVYIINOKADADKQVAFQVGVYVQKPNBGDLAKITGKTLTNGCVNQG
DANGNDLFEHRIYNDYGARIVNHNTHDITLTLGKSLITPNLSVSHIONDYED
DSYVYRPRRPIDQKDLKYKRYVALSGRGLNAPMPENGAENNDIMFQYTOBE
ARKNANMKNRRTIGDPGFEDEKGRNHLNENFKNQONRLLNGALNCKI
SVTQGNVLISGPTTHARDFVSKSRKRAHRSKNEVFEEDMIKRTKRAELIAYNO
SASFSSGRNVSDITANTITADAKVNLGKNGDEVCVRSDYTYVCNCGNISDKLN
SEFATRINGNVMNLQNALVILGKAALMGKIQOGNSRVSLNOKSHLITGDSYHNLS
LADSHIHNNASDAOSANKYHTIKIINHLGNGHFHYLTDLAKMLGDKVLVKEASGHY
OLHVNKTEPNOEGDILFEDASVODRSRLFVSLAHYVDLALRYTIKTEGITRLY
NPYANGPVPKPAAPANTASOAKATOTDQLOIAKPNIVYVAPSPONQAOEALR
QAKAEQYKROQAAREKVAROKDEAKKAKAELIARQEARAAELIARQEARAAE
ARELAKARAEASHOANAKPKRRRRALIPRAPYFSLDTPAKNSESICNILAR
IPRMGRRLINDYEETPLELEDEAEERQATQFHSKRRAISSPESDEASEBY
STSDKHPDONTLEHEKRYAGLOPRAOQRTQAAQADAVSTNLSLSDAMSTOSI
LIDGTGAYLRHTAOKSRADAERKSVMSNTGYRDYASAOYRFFSKRTQOIGIDRS
LSENNQIGCVLTVYSOHTFEDAGGKNTFVQVQALYKGYVINDMAYVADIGACSLRSR
LOTOQANFNRTSISQGTILGNTIKINOREIYVPSAGIRYSLSSADVKIDDSVKS
MAVKTLTGADIDPAVRKVNLTVPKPLISAIFYNRYGKGYNVGKSPATYAADNQOIS
AGVALYRNVLTNVNGSITRKQLEKSGQIKIQIRF"

BASE COUNT 1617 a 1105 c 1079 g 1098 t
ORIGIN

Query Match 4.5%; Score 195.6; DB 6; Length 4899;
Best Local Similarity 62.3%; Pred. No. 3.5e-33;
Matches 344; Conservative 0; Mismatches 199; Indels 9; Gaps 2;

1167 gaataatggaagaataatcttcttatttgcagaaggaagaaggaatgacttaccag 1226
1192 GAACAACGGACAAATGTTACCTT-----GAAGACACGGCAGCTTGGATTGATCA 1245
1227 caacatcaatcaagatgctgaggaattatattccaaagaggaatttcggtcgcctga 1286
1246 AATCATCAACCAAGGCGCGGCGCTGTTTTCAAAGGCGATTACACAGTCAAAAGCGC 1305
1287 aaataacgaa----acttgcagaagcgcgcttccatatacagtaagaacagtaaccgttac 1343
1306 AATATATGACATCATTGTTAGTGTGCGGGGATTGATGTCGCCAGCAAAAAGTCT 1365
1344 ttggaagaataaagcgcttgcagaacgacgcttgcagaatcgcgaaggaagcgctgca 1403
1366 TTGGCAAGTCAAAATCCGAATGCGCACAGATTGCGCAAAATCGCAAGGCACTTGTGA 1425
1404 cgttcaagccaaggggaaacccaagggctcgaacgcttgcagcgcttgcagcggtacagttat 1463
1426 AATTAACGGCACAGGCGGCTGTTAAAGGTGCGGCGGATGCGGATGCGTATGTTCT 1485
1464 ggaatcagcgagcagcagataaagcaaaaacgaagccttgcagtaacgcttgcag 1523
1486 GAATCAAAAAGCGGATGCCAGCAAAAAGTTCAAGCTTTCGCCAAGTCCGCGCATTTGTCAG 1545
1524 cggcaggggagctgctgaactgaatgcggaatacagttcaaccgcgaacactatlt 1583
1546 CGGACGGGATGATGATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1605
1584 cggcttcgagggaggaattggaattaaagggcagcttcgcttcgctccaggtattca 1643
1606 CGGTTTCGGGGGGGCTGTTGGATGCGCAATGGAATGACTTGTGACTTTTGAACACATCCG 1665
1644 aatatacgaagaagggagatgatttgcacaacacatacaagaagaagaatccacgcttca 1703


```
gene
<1..>2980
/gene="iga"
/allele="iga4"
<1..>2980
/gene="iga"
/codon_start=3
/transl_table=11
/product="IgA1 protease precursor"
/db_xref="GI:2529490"
/translation="LAVALTPESEALVVDVYDQJFRDFAENKGFYVADTSLVKN
KOGONIGNALSNVPMIDFSVADVNRLTVIDPOYAVSKHVGDEISYGHNGHLD
VSDNEENSVANNDYEPKNNHNGOGLIEDNMRINKFYEVPADIPATSGGVE
TYDKRSEFVAVGAGTOPEVNSRYNMEISRAITRATAGTYOVNTSINOGEL
IGFDNSKHSPEKLEKVSQNLTVAVLSDGSLFAYDKQEKRWPLGADVAG
YKNSQWENIYKKEFADIKORDNAGTIGYOEHEMKTTGTSHIGSTAVRAGNER
GANNGONVTFENNGLVLDONINOGAGGLFEKGDYVVKANNGITWLGADIVADCK
VWQVKNPNDRLAKIGKGTLENGTVNOGOLKVGDTVILNQADADKVOAFSQV
GIVSGGTVLNSPDQINPNLYEFGFRGLDANGDLFEHIRNDEGARIVNHNMTD
RASTITLTKSLITAPONSUYEIRNDYDDDDYGYYSRKPPOCKDLYKNRYYA
LKSQSVNAPMPENCGTENNDILMGSTOBEAKKNNMKNORISGSGFEENGK
GHGALINLNFNGKSAONRRLTGTNLNKGISTYOGCNVLISGRPTPHARDPVKSSAO
KHAHFSKNEVFEEDWIRNTRKATEITVNOSSAFSSGRVSNITANITATDAKVLN
GYKNGDEVCSRSDYGTVCNTGNSDKALNSFGATKINGVNLNQNALVILKALV
GOIOGNGSVSLNOSHKNLTDGSDVHNLSLADSHIHUNNASDAQSANOYHTLKINH
LSGNGHFHYLTHLAENLGDYLVKESASGHYOLHYODKTGEPOEGLNLFDSASVDR
SHLSVSLANHHVDLGLRYTIKTENGITRLYNPYAENRRRVKAPAPATNTASQAQTD
SAQIAKPONIIVVAPSPQAN"
```

BASE COUNT 974 a 687 c 642 g 677 t

Query Match 4.4%; Score 193; DB 1; Length 2980;
Best Local Similarity 57.4%; Pred. No. 1.2e-32;
Matches 368; Conservative 0; Mismatches 270; Indels 3; Gaps 1;

1198 gaaggaagaaggaattgacttaccagaacacatacaatgaagtgctggaagtattat 1257
1035 GAAGAACACGACCTTGTGATTCATCAACACCAACGCGCGCTCTCTT 1094
1258 ttcaagaagaatttcaggtcgcctgaataaaga---acttgaagaagcgcgagc 1314
1095 TTCAAGGCGCATTTACACAGTCAAGTGCACAAATATATGCGTCTGCTAGGCGAGG 1154
1315 gtccatacagtgaaagacagtlaccttacttgaagaataaagcggtgcaaacgacgc 1374
1155 ATTGATGTTGCCCGACGCAAAAAAGTCGTTGGCAAGTCAAAATCGCAATCGCACAGA 1214
1375 ctgtccaaatcggaagaagcagctgcagcttcaagcgaagaagggaaacgaagctcg 1434
1215 TTGGCAAAAATCGCAAGGCGACATTAGAAATTAACGCGCGCTTAACCAAGGCGCAA 1274
1435 atcagcggtggcgaagcagctacgtcatttggatcagcagaagcagaataaagcaaaa 1494
1275 TTAAAGTCGCGACGCTACGGTTATCTGAATCAACACCCGATCCGACAAAAAGTC 1334
1495 caagcctttagtgaataacgctgctgcaagcgaggggtacggtgcaactgaatccgat 1554
1335 CAGGCTTTCGCCAAGTGGCATTTGAGCGGCGCGCTACATTGTCGTGAACAGTCCG 1394
1555 aatcagttcaaccccgacaactctatcttcgcttcgcgcgagcagcttggatttaac 1614
1395 GACCAATTAACCCGAAATATCTGTATTGCTTCGTCGCGCTGTTGGATGCCAAT 1454
1615 gggccttcgcttccttccacglatcacaataacgaataaagggagatgattgtcaac 1674
1455 GGCAATGACTTGACTTTTGACACATCCGCACAGTGGATGAAGCGCGCCATTGTGAAC 1514
1675 caaatcaagaacaagaatcacacggttacatcagaagcaataaagatattgtctaacac 1734
1515 CACAACACAGACCGCGCTCCACATACATTAACGCGCAAAATCATTAATACCGCTCCT 1574
1735 ggcataataacaacgcttgatagcaaaaaagaatgtcctcaaacggttggttgagcag 1794

Db 1575 CAATATCTTCTGTTTATGAATCCGAATGATACATGATGACGATTATTATGCTAT 1634
Qy 1795 aaagatacagccaaagaacgagcgctcaactcttta 1835
Db 1635 TACAGCTATAGAAAACCATTCGCCAAGCGCAAGATCTTTA 1675

RESULT 33
AF012204 2980 bp DNA linear BCF 30-JAN-2001
LOCUS Neisseria meningitidis strain 23906 IgA1 protease precursor (iga)
DEFINITION gene, partial cds.
ACCESSION AF012204
VERSION AF012204.1 GI:2529491
KEYWORDS Neisseria meningitidis.
SOURCE Neisseria meningitidis.
ORGANISM Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;

REFERENCE 1 (bases 1 to 2980)
AUTHORS Morrell, G., Malorny, B., Muller, K., Seiler, A., Wang, J.-F., del Valle, J., and Achtman, M.
TITLE Clonal descent and microevolution of *Neisseria meningitidis* during 30 years of epidemic spread
JOURNAL MOL. MICROBIOL. 25 (6), 1047-1064 (1997)
MEDLINE 98010345
REFERENCE 2 (bases 1 to 2980)
AUTHORS Morrell, G., Malorny, B., Mueller, K., Seiler, A., Wang, J.-F., del Valle, J., and Achtman, M.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-1997) Max-Planck Institut fuer Molekulare Genetik, Inmestrasse 73, Berlin 14195, Germany
FEATURES location/Qualifiers
1..2980
/organism="Neisseria meningitidis"
/strain="23906"
/db_xref="taxon:487"
/note="serogroup A, subgroup III, isolated in China, 1966; also present in strain B1, serogroup A, subgroup IV-2 isolated in USA, 1937"
1..>2980
/gene="iga"
/product="IgA1 protease"

gene
CDS
mat_peptide
1..>2980
/gene="iga"
/product="IgA1 protease"
/protein_id="AAC45787.1"
/db_xref="GI:2529492"
/allele="iga2"
<1..>2980
/gene="iga"
/codon_start=3
/transl_table=11
/product="IgA1 protease precursor"

Db 1575 CAATATCTTCTGTTTATGAATCCGAATGATACATGATGACGATTATTATGCTAT 1634
Qy 1795 aaagatacagccaaagaacgagcgctcaactcttta 1835
Db 1635 TACAGCTATAGAAAACCATTCGCCAAGCGCAAGATCTTTA 1675

BASE COUNT 973 a 688 c 642 g 677 t

ORIGIN

Query Match 4.4%; Score 193; DB 1; Length 2980;
Best Local Similarity 57.4%; Pred. No. 1.2e-32;
Matches 368; Conservative 0; Mismatches 270; Indels 3; Gaps 1;

QY 1198 gaaggaagaagcgaattgacttaccagaacatacaatgaagtcgtgagattat 1257
Db 1035 GAAACACACGCACTTGGTATTTGATCAAAACATCACACGCGCGGCGTCTTT 1094
QY 1258 ttccagaagatttaccgttcgcctggaataacgaa---acttggcaagcgcg 1314
Db 1095 TTCAAAAGCGATTACACATCAAAAGTGCAAAATATGCGACTTGGCTAGGTCAGG 1154
QY 1315 gtcatatcagtaagaagacagtagtaccgttacttggaaagtaacggtgcaacgacgc 1374
Db 1155 ATTGATGTTGCCGACGCAAAAAGTCGTTGGCAAGTCAAAATCCGAATGGCGACAGA 1214
QY 1375 ctgtccaaatcggaagcgaagcgtgcacgttcgaagcgaaggggaaaccaaagctcg 1434
Db 1215 TTGGCAAAAATCGGCAAGGACATTAGAAATAAACGGACAGCGTTAACCAAGGCGAA 1274
QY 1435 atcagcgctggcgacgagtagacatcttggatcagcagcgacgaataaagcgaataa 1494
Db 1275 TTTAAAGTCGGCGACGCGTACGTTATTTCTGATCAACACAGCGGATGCCGACAAAAGTC 1334
QY 1495 caagccttagtgaatcggcttgcacgcgcaagggtaacgtgcaactgcaatgacgat 1554
Db 1335 CAGGCTTTCCTCCCAAGTCGGCATTTGAGCGCGCGGCGGCTGATGTCGACAGTCGC 1394
QY 1555 aatcattcaaccccgcaaacacttattcggttcgcgacgagcgttggattaaac 1614
Db 1395 GACCAATTTAACCCGATTAATGATATTCGTTCCGTGCGGCTGTTGATGCCAAT 1454
QY 1615 gggcattcgcttcgttccacgcttaccataaataacgaataagggcgatgattgcaac 1674
Db 1455 GCGAATGACTTGAATTTTGAACACATCCGACAGCGATGAGGCGCGCATTTGTCAAC 1514
QY 1675 caaatcaagaacaaagaatccacgcttaccatcgaagcaataagattatgataaac 1734
Db 1515 CACACACACAGACGCGCGCTCCACATCACTTAACCGGCAATCTTAAATTAACCGCTCT 1574
QY 1735 ggcataacacacagcttgcgtagacaataaagaattgcctcaacggttggcgag 1794
Db 1575 CAAAATTTCTGTTTATGAAATCCGAATGATTAAGATGATGAGATTAATATGCGTAT 1634
QY 1795 aaagatacgaacaaacgaacgagcggtcgaactgttta 1835
Db 1635 TACAGCTATAGAAACCATTCCGCAAGGCAAGATCTTTA 1675

RESULT 34
AF012205 2980 bp DNA linear BCT 30-JAN-2001
LOCUS Neisseria meningitidis strain Z3910 1gal1 protease precursor (19a)
DEFINITION gene, partial cds.
ACCESSION AF012205
VERSION AF012205.1 GI:2529493
KEYWORDS
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis; beta subdivision; Neisseriaceae; Neisseria.

REFERENCE
AUTHORS Morelli, G., Malorny, B., Muller, K., Seiler, A., Wang, J. F., del
TITLE Vallée, J. and Achtman, M.
JOURNAL Clonal descent and microevolution of Neisseria meningitidis during
MEDLINE 30 years of epidemic spread
REFERENCE Mol. Microbiol. 25 (6), 1047-1064 (1997)
AUTHORS 2 (bases 1 to 2980)
Morelli, G., Malorny, B., Mueller, K., Seiler, A., Wang, J. F., del

Vallée, J. and Achtman, M.

TITLE Direct Submission
JOURNAL Submitted (02-JUL-1997) Max-Planck Institut fuer Molekulare
Genetik, Inmestrasse 73, Berlin 14195, GermanyFEATURES
source Location/Qualifiers

1. .2980

/organism="Neisseria meningitidis"

/strain="Z3910"

/db_xref="taxon:487"

/note="serogroup A, subgroup III; isolated in China, 1966"

1. .>2980

/gene="iga"

/product="Igal1 protease"

<1. .>2980

/gene="iga"

/allele="igal1"

<1. .>2980

/gene="iga"

/codon_start=3

/transl_table=11

/product="Igal1 protease precursor"

/protein_id="AAC45788.1"

/db_xref="GI:2529494"

/translation="LATATPTYSALVDDVDYQIFRDAEMKGFPGATPLSVN
KOGONIGALSNVPMIDESVADVNRTLVVDPOAVSVKHKGDISTYGHNHID
VSNDENERSVANDYEPNKNMHGNGLEENYMARLKEVTEVAPLAPTSAGGVE
TYKDNRESEFVRAAGTOFEYNSRNPTELRAVYALAGPEYDQVNTVSNLQEG
IGFDNKHSHLEKLEVLSONALTNVAVLIGDSGPLFAVDKESHEVIGADVANG
YOKNMOENIYIKERPADEIKORDNAGTIGGHEHMKTTGTSGSTVRAAGNER
GANNCONVTFENNGTIVLDONINOGAGLFPFGDYVKNANGITMAGIDVADKK
VVOVKNPNDRLAKIGKSTLEINGVNOGOLKVEDGYIILNOADADKKVOAFOY
GIVSGRGLVLNLPDQINPNNTYFGRGRLDNGNDLVEHIRNVDGARLYNHNTD
RASTITLTKESLITAPONLSVEIRNDYDDDDYGYYSRKLPIQGDILYKNRTYA
LKSQSVNAPMECOTENNMDILMSTOEAKNMMNHKNORISGSGFEEENCK
GHNALNLNENKSAQNRFLTGTNLNGKISYVQGNVLISGPTPHARPVFNSSAO
KDAHKSNNVEYFEDDMINRTFATETLVWOSAFSSGRNVSNTATITADNKAHVL
GYNKDEVCVRSRDYTYGVCNCGNISDKALNSGAGQIINCNVILNOMAAVLGLKALM
GQIQGGSNRSLSNOKSHMLTSDOSVHNLSLADSHIHLNNAADGSAOANYHTLKIH
LSQNGHFHYLTHLAENLGDKVLVKEASGHYQLHVQDKTEPNOEGNLNLDASSVR
SHLSYSLANHHVDLGLARYTIKTENGITRLYNPYAENRRRVKPPAPATVNTASQOTD
SAQIAKPQNVIVAPPSPOAN"

BASE COUNT

973 a

687 c

642 g

678 t

ORIGIN

Query Match 4.4%; Score 193; DB 1; Length 2980;
Best Local Similarity 57.4%; Pred. No. 1.2e-32;
Matches 368; Conservative 0; Mismatches 270; Indels 3; Gaps 1;

QY 1198 gaaggaagaagcgaattgacttaccagaacatacaatgaagtcgtgagattat 1257
Db 1035 GAAACACACGCACTTGGTATTTGATCAAAACATCACACGCGCGGCGTCTTT 1094
QY 1258 ttccagaagatttaccgttcgcctggaataacgaa---acttggcaagcgcg 1314
Db 1095 TTCAAAAGCGATTACACATCAAAAGTGCAAAATATGCGACTTGGCTAGGTCAGG 1154
QY 1315 gtcatatcagtaagaagacagtagtaccgttacttggaaagtaacggtgcaacgacgc 1374
Db 1155 ATTGATGTTGCCGACGCAAAAAGTCGTTGGCAAGTCAAAATCCGAATGGCGACAGA 1214
QY 1375 ctgtccaaatcggaagcgaagcgtgcacgttcgaagcgaaggggaaaccaaagctcg 1434
Db 1215 TTGGCAAAAATCGGCAAGGACATTAGAAATAAACGGACAGCGGTTAAACCAAGGCGAA 1274
QY 1435 atcagcgctggcgacgagtagacatcttggatcagcagcgacgaataaagcgaataa 1494
Db 1275 TTTAAAGTCGGCGACGCGTACGTTATTTCTGATCAACACAGCGGATGCCGACAAAAGTC 1334
QY 1495 caagccttagtgaatcggcttgcacgcgcaagggtaacgtgcaactgcaatgacgat 1554
Db 1335 CAGGCTTTCCTCCCAAGTCGCGATTCGAGCGCGCGGCTGATGTCGACAGTCGC 1394

[illegible][illegible]

YINDAWYVAGDIGAGSLRSLQOQKANKFNRTSIOTGLTIGNTLKIINOPIEVSAGIR
YSLRSADYKLGINDSVKSMYKTLTAGIDFAYRRKVGITPKPLISAAYFANYKKG
GVNNGNSFAYKADNOQXYSAGAAALIXRVTLNVNSITRKQLEKOKSQTIKIQIRE
"

BASE COUNT 1542 a 1074 c 1041 g 1002 t

ORIGIN

Query Match 4.4%; Score 193; DB 1; Length 4659;
Best Local Similarity 62.0%; Pred. No. 1,3e-32;
Matches 333; Conservative 0; Mismatches 195; Indels 3; Gaps 1;

QY 1198 gaaggaagaagaatgtaacttaccgaacaatcaatcaatgctgagagattatc 1257
DB 1084 GAAGGACACGCGACTTGTGATTGGATCAAAACATCAACGCGCGGCGCTCTTT 1143
QY 1258 ttccaaggaatttcagctcgcctgaataaagaa---acttgcaagcgcgagc 1314
DB 1144 TTCAAGGGGATTACACAGTCAAAAGTATTAATATGACATCTGCTAGTGGCGGG 1203
QY 1315 gtccatacagtgaaagacagctaccgttaacttggaagtaacgagcgctgcaacgaccgc 1374
DB 1204 ATTGATGTTATCCAGCGCAAAAAAGTCTTGGCAACTCAAAATCCGAACGCGCAGCA 1263
QY 1375 ctgtccaaaatcgcaaaagcagctgcacgttcaagcgaaggaaggaagcagctcg 1424
DB 1264 TTGGCAAAATCGGCAAAAGGCACATTAGAAATTAACGGCACAGCGCTTAAACAAAGGCAA 1323
QY 1435 atcagcgctgggagcaggtacagtcatttgatcgaagcagcagcgaataagcaaaaa 1494
DB 1324 TTTAAAGTCGCGCAGCGTACGTTATCTGAATCAACACCGATGCGCAAAAAAGTGC 1383
QY 1495 caagccttgatgaatcgctgctgacgagcgaggtacgctgcaactgaatcgccgat 1554
DB 1384 CAGGCTTTCCCAAGTGGCGCATTTGTACGGGACGCGGTACATTGTTAATATAGTTCA 1443
QY 1555 aatcagttcaaccgcgaacaatctatcttcgcttcgcgagcgagcttgagattaac 1614
DB 1444 AATCAATTAATATCCGATTAATATTTCCGTTTCCGTCGCTGCTTGGATGCCAAT 1503
QY 1615 ggcattcgcttcgctcccgatcacaataacgaatgaagcgagcgatgctgaac 1674
DB 1504 GGCAATGACTTGACTTTTGAACACATCCGCAACGTGATGAAGCGCGCGCATTTGTCAAC 1563
QY 1675 cacatcaagaagaagaatccaccgttacattacaggaac 1715
DB 1564 CACAACAACAAGCCAGCGCTCCACGATCACGCTAACGGGTAA 1604

RESULT 38
NMA322491/c 311321 bp DNA linear BCT 04-DEC-2000
LOCUS
DEFINITION Neisseria meningitidis serogroup A strain Z2491 complete genome;
segment 3/7.
ACCESSION AL162754 AL157959
VERSION AL162754.2 GI:7379424
KEYWORDS
SOURCE
ORGANISM Neisseria meningitidis Z2491.
Bacteria; Proteobacteria; Delta subdivision; Neisseriaceae;
Neisseria.

REFERENCE 1 (bases 1 to 311321)
AUTHORS Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C., Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T., Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N., Holtroyd,S., Jagsels,K., Leather,S., Moule,S., Mungall,K., Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M., Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
TITLE Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
JOURNAL Nature 404 (6777), 502-506 (2000)
MEDLINE 20222556
REFERENCE 2 (bases 1 to 311321)

AUTHORS TITLE JOURNAL

COMMENT

Parkhill,J.
Direct Submission
Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
Sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
Notes:
Details of N. meningitidis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).

FEATURES

source

1..311321 "Neisseria meningitidis Z2491"

gene

/db_xref="taxon:122587"

CDS

/note="serogroup: A"

misc_feature

/note="NMA0698"

gene

/note="NMA0698", possible periplasmic protein, len: 152 aa;
unknown, contains a probable N-terminal signal sequence
and lies within a region of unusually low GC content"

CDS

/codon_start=1
/transl_table=11
/product="putative periplasmic protein"

gene

/db_xref="GI:7379425"

CDS

/db_xref="SPTREMBL:O9JY02"

gene

/translation="MLTKLILFLFVFLVLAINLFFESSDIESFGNYOFEYVD
KCPANVILVMKDNENFDKIIISGLVEYKDDNIFYEYIDGGVSAOSCYYKPEI
LYGKIILNKHIINISMENKNEFLSEDKIMKGTNNMLADPKNKNQITLD"

CDS

/note="NMA0699"

gene

/note="NMA0699", pseudogene, probable ABC transporter
protein, len: 336 bp; similar to C-terminal of many ABC
transporters e.g. TR:Q46973 (EMBL:U47048), mtfB,
Escherichia coli microcin transport protein (707 aa),
fasta scores: E(): 2.6e-19, 53.2% identity in 111 aa
overlap. Contains PS00211 ABC transporters family
signature. NMA0686 may be the remainder of this
pseudogene"

CDS

/codon_start=1
/pseudo
/transl_table=11
/product="putative ABC transporter protein (pseudogene)"

gene

/note="NMA0699"

CDS

/note="PS00211 ABC transporters family signature"

gene

/note="NMA0700"

CDS

/note="NMA0700"

misc_feature

/note="NMA0700, possible ribonuclease BN, len: 408 aa;
similar to 5W:RBN_ECOLI (EMBL:L19201), rbn, Escherichia
coli ribonuclease BN (EC 3.1.1.-), fasta scores: E():
2.3e-28, 34.3% identity in 266 aa overlap. Longer than rbn
at the C-terminus. Also similar to TR:O85449
(EMBL:AF067083) Vitreoscilla sp. hypothetical protein (376
aa), fasta scores: E(): 0, 45.0% identity in 333 aa
overlap. Contains a region similar to NMA0677, fasta
scores: E(): 4.3e-08, 64.4% identity in 45 aa overlap.
Contains hydrophobic, possible membrane-spanning regions,
as does rbn"

gene

/codon_start=1
/transl_table=11
/product="putative ribonuclease BN"

CDS

/protein_id="CA83987.1"

misc_feature

/db_xref="SPTREMBL:O9JY79"

translation

"MTFLORLGLADNKCICAFWYVRRFDEERVPDAASMTFTTL

OY 1735 ggcataacacacgcttgatagcaaaagattgctacacagcttgctgagc 1794
 Db 188360 CAAAATCTTCTGTTATGAATCCGAATGATATGACGATTTATGGCTAT 188301
 OY 1795 aaagatacgaacaaacgaagcgctacacctgttta 1835
 Db 188300 TACAGCTATAGAAACCATTCCGCAAGCGCAAGATCTTTA 188260

RESULT 39

NMIGAC65
 LOCUS NMIGAC65 1525 bp DNA linear BCT 14-MAR-1995
 DEFINITION N.meningitidis iga gene (NGC65).
 ACCESSION X82484
 VERSION X82484.1 GI:732871
 KEYWORDS iga gene; IgaI protease.
 SOURCE Neisseria meningitidis.
 ORGANISM Neisseria meningitidis
 Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;

REFERENCE 1 (bases 1 to 1525)
 Lomholt,H., Poulsen,K. and Kilian,M.
 Comparative characterization of the iga gene encoding IgaI protease
 in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
 influenzae
 JOURNAL Mol. Microbiol. 15 (3), 495-506 (1995)
 MEDLINE 95302961
 AUTHORS Lomholt,H.
 TITLE Direct Submission
 JOURNAL Submitted (19-OCT-1994) H. Lomholt, Dept. of Medical Microbiology,
 The Bartholin Building, University of Aarhus, 8000 C Aarhus C,
 DENMARK

FEATURES

source 1.1525 Location/Qualifiers
 /organism="Neisseria meningitidis"
 /strain="NGC65"
 /db_xref="taxon:487"
 1.1525
 /gene="iga"
 <1..>1525
 /gene="iga"
 /codon_start=1
 /transl_table=11
 /product="IgaI protease"
 /protein_id="CA57867.1"
 /db_xref="GI:732872"
 /db_xref="SPTREMBL:O51168"
 /translation="ATDLSYKKGKOGNIGNALSNVMDISVADVNRATVYVDPYA
 VSKHAKAEVHTFYGYOVYNGHNDVADKENEYVVDNNYEPHKASASLGRLEEDNM
 ARFNKFEVAVIAPFDAGGGLDITKDKRESFVGVGROLVTEKGAHOBGNEKG
 YDLRLDSQAVRYAIAGTPIKIDINIDOTMTBESLIGGNINITYSAEELQALSDALT
 NGVLGDSGSPLEAFDKONQWFLGTQYVAGYKRSQWENYIKKEFKADIKQDN
 AGTIGNGEHNNITFTGNSHIGSTAVRIAGMEKANNQNTYFEDNGTLVDQINQ
 GAGGLEFKDGYVKGANNIDITWLAGIDVADKRVQWYVSGRPLDLAGTIGTLEIN
 GTGVQGGKLVGDTVILNOKADSNOKVAFSGVIGSRRLVLVNSSQINDPDLNF
 GPRGRDLANGNDLTFEHRNDEGARIVNHTGASTITLIGKSLITRAPQNLSTYEI
 "

BASE COUNT 510 a 320 c 330 g 365 t
 ORIGIN

Query Match 4.4%; Score 192.4; DB 1; Length 1525;
 Best Local Similarity 60.8%; Pred. No. 1.5e-32;
 Matches 332; Conservative 0; Mismatches 211; Indels 3; Gaps 1;

OY 1198 gaagaaagcgaaattgatactaccagcaacatcaatcaagctgtaggagatat 1257
 Db 961 GAAGACAAGCGCACTTGGATTGGATCAAAACATCAACCAAGCGCGGCTGTTT 1020
 OY 1258 ttccaagagatttcgctgcctcgctgaataaagaa---acttggcaagcgcgggc 1314
 Db 1021 TTCAAAGGCGATTACACGTCAAGGTGCAATTAATGACATCTGCTAGGTGCGGG 1080

OY 1315 gtccatacgtgaagaacagctaccgttacttggaagaaagcgcgtgcaaacgcgc 1374
 Db 1081 ATTGATCTTCCCGCAGCGCAAAAAGTCGTTGGCACTCAAAAATCCGAATGGTACAGA 1140
 OY 1375 cgtccaaatcgtgcaaaagcgtgacgcttcaagccaaaggggaaacaaagctcg 1434
 Db 1141 TTGGCAAAAATCGGCAAGCGCATATTGAATAAACCGCAGCGCTTAACCAAGGCGCA 1200
 OY 1435 atcagcgtggcgagcagcttaccatcttgatccagcagcagcagcagcagc 1494
 Db 1201 TTTAAATCGCGCGACGACAGATTATCTGATCAATCAAAAACCGATTCAAAAAGTC 1260
 OY 1495 caagccttagtgaatcgcgttgctcagcgcaaggggtacgylgcaactgtaagcgc 1554
 Db 1261 CAGGCTTCTCCCAAGCGGCGCATGTCAGCGGCGCTACATTGGTATTAATAGTTCA 1320
 OY 1555 aatcaatcaaccccgcaacacacacacacacacacacacacacacacacacac 1614
 Db 1321 AATCAGATTATCCGATTAACCTATATTTCGCTTCGCTGCGGCTCTTTGGATGCCAAT 1380
 OY 1615 gggcatcgcttcgttcacacgctatcaaaatacagcagcagcagcagcagc 1674
 Db 1381 GGCAATGACTTGACTTTTGAACACATCCGACAGCGTGATGAAGCGCGCATGTCAAC 1440
 OY 1675 caaatcaagaacaaagatccacgcttaccattacagcagcagcagcagcagc 1734
 Db 1441 CACACACAGCGCCACGCTCCACATCACTTAACCGGCAATCAATTAATTCACGCTCT 1500
 OY 1735 ggcact 1740
 Db 1501 CAAAAT 1506

RESULT 40

NMIGAL39
 LOCUS NMIGAL39 1489 bp DNA linear BCT 14-MAR-1995
 DEFINITION N.meningitidis iga gene (81139).
 ACCESSION X82477
 VERSION X82477.1 GI:732856
 KEYWORDS iga gene; IgaI protease.
 SOURCE Neisseria meningitidis.
 ORGANISM Neisseria meningitidis
 Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;

REFERENCE 1 (bases 1 to 1489)
 Lomholt,H., Poulsen,K. and Kilian,M.
 Comparative characterization of the iga gene encoding IgaI protease
 in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
 influenzae
 JOURNAL Mol. Microbiol. 15 (3), 495-506 (1995)
 MEDLINE 95302961
 AUTHORS Lomholt,H.
 TITLE Direct Submission
 JOURNAL Submitted (19-OCT-1994) H. Lomholt, Dept. of Medical Microbiology,
 The Bartholin Building, University of Aarhus, 8000 C Aarhus C,
 DENMARK

FEATURES

source 1.1489 Location/Qualifiers
 /organism="Neisseria meningitidis"
 /strain="81139"
 /db_xref="taxon:487"
 1.1489
 /gene="iga"
 <1..>1489
 /gene="iga"
 /codon_start=1
 /transl_table=11
 /product="IgaI protease"
 /protein_id="CA57860.1"
 /db_xref="GI:732857"
 /db_xref="SPTREMBL:O51163"

FEATURES	source
gene	
LOCUS	NMIGA341
DEFINITION	N.meningitidis iga gene (IHN5341).
VERSION	X82478.1
KEYWORDS	iga gene; Iga1 protease.
SOURCE	Neisseria meningitidis.
ORGANISM	Neisseria meningitidis Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
REFERENCE	1 (bases 1 to 1489)
AUTHORS	Lombolt,H., Poulsen,K. and Killian,M.
TITLE	Comparative characterization of the iga gene encoding Iga1 protease in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus influenzae
JOURNAL	Mol. Microbiol. 15 (3), 495-506 (1995)
MEDLINE	95302961
REFERENCE	2 (bases 1 to 1489)
AUTHORS	Lombolt,H.
TITLE	Direct Submission
JOURNAL	Submitted (19-OCT-1994) H. Lombolt, Dept. of Medical Microbiology, The Bartholin Building, University of Aarhus, 8000 C Aarhus C, DENMARK
FEATURES	Location/Qualifiers
source	1. 1489
	/organism="Neisseria meningitidis"
	/strain="IHN5341"
	/db_xref="taxon:487"
	1. 1489

/product="Iga1 protease"
/db_xref="PID:el23968"
BASE COUNT 498 a 319 c 321 g 351 t
ORIGIN

Query Match 4.3%; Score 189.8; DB 1; Length 1489;
Best Local Similarity 61.6%; Pred. No. 5.8e-32;
Matches 321; Conservative 0; Mismatches 197; Indels 3; Gaps 1;

QY 1198 gaagaaagcgaattgacttaccgaacatcaatcaatgctgagattat 1257
DB 925 GAAACACAGGACCTTGGATTGATCAAAACATCAACAGCGCGGCTGT 984
QY 1258 ttccaaagatcttaccgctcgcctgaaataacgaa---acttggcaagcgcgcc 1314
DB 985 TTCAAAGGCGATTACACACTCAAAAGTATTAATATGACATCACTTGGCTAGTGGCGG 1044
QY 1315 gtccatatacgtgaagacgcttaccgttgaagaataacggcggtgcaaacgcgc 1374
DB 1045 ATTGATGTTGCGGACGCGGCAAAAGTCTTTGGCAAGTCAGAAATCCGAATGGCGACAGA 1104
QY 1375 ctgtccaaatcgcaagcgcgcctgcacgltcaagcgaaggggaaacgaagctcg 1434
DB 1105 TTGGCAAAATCGGCAAGGACATTAAGAAATTAACGGCAGCGCTTAACCAAGGCGCA 1164
QY 1435 atcagcgltggcgagcggtacgttatttgatcagcagcgacgataaaggcaaaaa 1494
DB 1165 TTAAAGTGGGCGGCGGTATTCGATTCGAACCAAGGCCATCCGCAAAAAGTCT 1224
QY 1495 caagccttagtgaagcgttgcagcgagcgaggtgcaactgaatgcgcgt 1554
DB 1225 TAGGCTTTCCCAAGTGGCAATGTCAGCGGACGGGTAACATGTTATTAATATGTC 1284
QY 1555 aatcaggttaaccccgcaaacactctatctcgcttcgcggcgagcggttgaattaac 1614
DB 1285 AATCAATTAATCCGATTAACCTATTAATTCGTTCCGTCGCGGTGGTGGACCAAT 1344
QY 1615 gggcattcgcttcgttccacggtatcaaaataacgataagggcgatgattgcaac 1674
DB 1345 GGCATGATTCGACTTTTGAACACATCCGCAACGTGATGAAGCGCGCATTTGTCAC 1404
QY 1675 cacaatcaagaagaatccacggttaccattacagca 1715
DB 1405 CACACACAGCGCACGCTTCACAAATCAGCGTAACGGGTAA 1445

RESULT 44
NMIGAF80 1489 bp DNA linear BCT 14-MAR-1995
LOCUS N.meningitidis iga gene (NC880).
DEFINITION X82479.1 GI:732883
VERSION X82479.1 GI:732883
KEYWORDS Iga gene; Iga1 protease.
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.

REFERENCE 1 (bases 1 to 1489)
AUTHORS Lomholt,H., Poulsen,K. and Kilian,M.
TITLE Comparative characterization of the iga gene encoding Iga1 protease
in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
influenzae
JOURNAL Mol. Microbiol. 15 (3), 495-506 (1995)
MEDLINE 95302961
REFERENCE 2 (bases 1 to 1489)
AUTHORS Lomholt,H.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-1994) H. Lomholt, Dept. of Medical Microbiology,
The Bartholin Building, University of Aarhus, 8000 C Aarhus C,
DENMARK

FEATURES
Source Location/Qualifiers
1..1489

/organism="Neisseria meningitidis"
/strain="NC880"
/db_xref="taxon:487"
1..1489
/gene="iga"
/gene="iga"
CDS
<1..>1489
/codon_start=1
/transl_table=11
/product="Iga1 protease"
/protein_id="CA57862.1"
/db_xref="GI:732884"
/db_xref="SPTREMBL:Q51172"
/translation="ATDLSVKNKQCONIGALSNNPMDISVADVNRRTLVYIDPOYA
VSKHVKDEISYUYNHNGHLDVSNDENYRSVAONDYEPNKNMHNGOGLIEDY NMA
RLNKFVTEVAPLAPTSAGGVEYTKDNFSEFVAGACTOFEYNSRYNMTELSRAY
VALAGPQDVAVTSLNDEGLIGEDNSKHSPELKEVLSONLNTYAVLGDGSP
LFAVXKPEHWFELGADYDYWAGYOKSKSDEMYLKKPEADEIKORPNAGTIGNENH
WTTGNSHIGSTAVRLANRERDANNGOVTEENNCTLYLDONINGAGGLPEKDYT
VGINDDITWLAGIDVADGKRVWVKPNQDRILKIGKLTLEINGTVNOSOLKVG
DQTVLINDQADADKRVQAFSVOYGVISGRRLTVLSSNOINPNLXFGFGRGLDANGN
DLTFEHRVWDEGARLVNHNTHASITLTGLKSLITDPTISIH1"

BASE COUNT 498 a 321 c 320 g 350 t
ORIGIN

Query Match 4.3%; Score 189.8; DB 1; Length 1489;
Best Local Similarity 61.6%; Pred. No. 5.8e-32;
Matches 321; Conservative 0; Mismatches 197; Indels 3; Gaps 1;

QY 1198 gaagaaagcgaattgacttaccgaacatcaatcaatgctgagattat 1257
DB 925 GAAACACAGGACCTTGGATTGATCAAAACATCAACAGCGCGGCTGT 984
QY 1258 ttccaaagatcttaccgctcgcctgaaataacgaa---acttggcaagcgcgcc 1314
DB 985 TTCAAAGGCGATTACACACTCAAAAGTATTAATATGACATCACTTGGCTAGTGGCGG 1044
QY 1315 gtccatatacgtgaagacgcttaccgttgaagaataacggcggtgcaaacgcgc 1374
DB 1045 ATTGATGTTGCGGACGCGGCAAAAGTCTTTGGCAAGTCAGAAATCCGAATGGCGACAGA 1104
QY 1435 atcagcgltggcgagcggtacgttatttgatcagcagcgacgataaaggcaaaaa 1494
DB 1165 TTAAAGTGGGCGGCGGTATTCGATTCGAACCAAGGCCATCCGCAAAAAGTCT 1224
QY 1495 caagccttagtgaagcgttgcagcgagcgaggtgcaactgaatgcgcgt 1554
DB 1225 TAGGCTTTCCCAAGTGGCAATGTCAGCGGACGGGTAACATGTTATTAATATGTC 1284
QY 1555 aatcaggttaaccccgcaaacactctatctcgcttcgcggcgagcggttgaattaac 1614
DB 1285 AATCAATTAATCCGATTAACCTATTAATTCGTTCCGTCGCGGTGGTGGACCAAT 1344
QY 1615 gggcattcgcttcgttccacggtatcaaaataacgataagggcgatgattgcaac 1674
DB 1345 GGCATGATTCGACTTTTGAACACATCCGCAACGTGATGAAGCGCGCATTTGTCAC 1404
QY 1675 cacaatcaagaagaatccacggttaccattacagca 1715
DB 1405 CACACACAGCGCACGCTTCACAAATCAGCGTAACGGGTAA 1445

RESULT 45
NMIGAF48 1489 bp DNA linear BCT 14-MAR-1995
LOCUS N.meningitidis iga gene (HF48).
DEFINITION X82480.1 GI:732875
VERSION X82480.1 GI:732875

KEYWORDS
SOURCE
ORGANISM

iga gene: Iga1 protease.
Neisseria meningitidis.
Neisseria meningitidis.
Bacteria: Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 1489)
Lomholt, H., Poulsen, K. and Kilian, M.
Comparative characterization of the iga gene encoding Iga1 protease
in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
influenzae

JOURNAL
MEDLINE
REFERENCE

Mol. Microbiol. 15 (3), 495-506 (1995)
95302961
2 (bases 1 to 1489)
Lomholt, H.
Direct Submission
Submitted (19-OCT-1994) H. Lomholt, Dept. of Medical Microbiology,
The Bartholin Building, University of Aarhus, 8000 C Aarhus C,
DENMARK

FEATURES

source

Location/Qualifiers
1. 1489
/organism="Neisseria meningitidis"
/strain="Hf48"
/db_xref="taxon:487"

gene

1. 1489
/gene="iga"
<1..>1489
/gene="iga"

CDS

/codon_start=1
/product="Iga1 protease"
/protein_id="CA57863.1"
/db_xref="GI:732876"
/translation="ARDLSVKKKQGNIGNALSNVPMIDPSVADVNRRLTYIDPOYA
VSKRYKDEISLYGHNGHLDVSDNETRYSYQNDIEPKNMHNGNORLSDYDMA
RLNKFVTEVAPIAPISAGGVELEYKDNKRFSEVRVAGTQFEYNSRYNMTLSRAVR
VAIAGTPYQDVVNTSNLQNGELIGFDNSKHSPEKLEVLSONALJNYAVLDSGSP
LEAYDKERRWVFLGAYDWAYGOKNSQWEMNLYKKRFADDEIKQDNACTIKGYGEHH
WRTYTNSHIGSTAVRLAGNERKANGOUNFENNGLTVLDQINOGAGLFFPGDYT
VGINDITWLGAGIDVADGKRVWQVKNPNNGRLKIGKLTLEINGTGVNOGOLKVG
DGVTLINOKADSNQVQAFQSOYGIYVSGRTLVINSNOINPDMLYFGRGRLDANGN
DLTFEHIRVDCARIVNHNHTGTSITTLTGKSLITNPNSLSVHSI"

BASE COUNT
ORIGIN

496 a 319 c 325 g 349 t

Query Match

Best Local Similarity 4.3%; Score 188.2; DB 1; Length 1489;
Matches 320; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

QY 1198 gaaggaagaagcgaattgatactaccagcaatcatcaatgaagtgtgtgagattatat 1257
DB 925 GAAACACACGCGACTTGTGATTGGATCAAAACATCAACCAAGCGCGGCTGTGTTT 984
QY 1258 ttcaagagagatttaccgctcgcctgaaataacgaa--acttgcgaagcgcgagc 1314
DB 985 TTCAAAGCGCATTAACAGCAATGAATGAATGATGATCACTTGTGCTAGTGCAGG 1044
QY 1315 gtccataatcagtaagacatgacttacttgaataaagcgcgtgcgaacgacgcg 1374
DB 1045 ATTGATGTTCCCGACGCAAAAAGTCTTGGCAAGTCAAAAATCCGAATGGGACAGA 1104
QY 1375 cgtlccaaatcgcgaagcagcgtlcaacgtlcaagccaagggaaggaacccaagctc 1434
DB 1105 TTGGCGAAATCGCAAGGCGACATTAAGAAATTAACGACAGCGTTAAACCAAGGCAA 1164
QY 1435 atcagcgtggcgaagcgtacagtcatttgcatttcagcagcagcagcagcagcagc 1494
DB 1165 TTTAAAGTGGCGACGCTAGCTATTCTGAATCAAAAAGCCGTTCAAAACCAAAAAGTC 1224
QY 1495 caagccttagtgaatcgcgttgcacgcgcagcgaggtacgcgtlcaactgaatgcgat 1554
DB 1225 CAGGCTTTCCTCCCAAGTCGCGCATTTGTCAGCGGACGCGTACATTTGTAATTAATGTTCA 1284

QY 1555 aatcagttcaaccccgacaacactatattgcttcgcggcggaagcttggatttaaac 1614
DB 1285 AATCAGATTATATCCCATTAACCTATATTGCTTCGTCGCGCGCTCGTTGATGCCAAT 1344
QY 1615 gggcattcgttcgttccacgcgtatccaataacgatgaaggcgcatgatgtcaaac 1674
DB 1345 GCGAATGACTTGAATTTTGAACATCCGCAACGCTGATGAGCGCGCGCATTTGTCAC 1404
QY 1675 cacaatcagacaagaatccaccgtttaccattacagga 1715
DB 1405 CACACACACAGCCACACCTTCACAAATCACGCTAACGGGTA 1445

Search completed: June 30, 2002, 11:42:53
Job time: 66391 sec

Db 1261 AACACGAAAGCTGGCAGAGCGCGGCTTCATATCATGTAGACAGTACCCTTACTTGG 1320
Oy 1321 aagaataacgctgagcaaaacgacgcctgtccaaaatcggcaaaagcagctgcagct 1380
Db 1321 AAGTAAGCGCTGGCAAAAGCAGACCGCTGTCCAAAATTCGCAAAAGGACGCTGCAGCTT 1380
Oy 1381 caagccaaggggaaaaacaaagctcgatcaagcgtggcgagcgttacagtaatttgat 1440
Db 1381 CAAGCCAAAGGGGAAAAACCAAGGCTCGATCAGCTGGCGAGCGGTACAGTATTTGGAT 1440
Oy 1441 cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1500
Db 1441 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1500
Oy 1501 aagggtacgctgcaactgaaatgcgcataatcagttcaaccccgacaacatctatttcggc 1560
Db 1501 AGGGGTACGGTGCACATGAAATGCGGATATCAGTTCAACCCGCAACCTCATTTTCGGC 1560
Oy 1561 ttccgcgcgagcagcttggatttaaacgggcatcgccttgccttcacgcgattataaat 1620
Db 1561 TTTCCGGCGCGAGCTTTGGATTTAACGGGCAATTCGCTTGTCCACCGTATTTCAAAAT 1620
Oy 1621 accgatgaaggcgatgatctgncatcataatgccaacaacatccacgctiaacat 1680
Db 1621 ACCGATGAAGGCGGATGATTGNCNATCATATGCGCACACATCCACCGTTACAT 1680
Oy 1681 acagggaatgaagtatatacaaaccgagtggtlaagaatataagacttaattacagc 1740
Db 1681 ACAGGGAATGAAGTATTTACACAAACCGAGTGAAGATATCAATATGACTTATTTACAC 1740
Oy 1741 aaagaaatctgctcaaacggtgtgtttggcgagaaagatacagcacaacagcagcgcg 1800
Db 1741 AAAGAAATTTGCTCAACAGGTTGGTTGGCGAGAAATATACACCAACCAACGAGCGCG 1800
Oy 1801 ctaaccctgttaccagcccgcaagaagacgcacccngcttgcttcgcgcgagaca 1860
Db 1801 CTCAACTTGTTTACGAGCCCGCGCAGAGACCGCACCCNGCTGCTTCCGGCGAGACA 1860
Oy 1861 aatttaaacgcaacatcagcaaaacagcgaactgttttaagcgcgagccgaca 1920
Db 1861 AATTTAACCGCAACATCACCAACAAACGCAAACTGTTTTTCAGCGCGAGACCGACA 1920
Oy 1921 ccgacgcctcaaatcatltaggaagcggtgtcaaaaatgaaagtatccacacaaga 1980
Db 1921 CCGACGCTCAATCATTTAGGAAGCGCGGTGTCAAAAATGGAAGATTCACCAAGGA 1980
Oy 1981 gaatcgtgtggagacaacgactgatacnaacgcgacgcttaaaacggaataatccatc 2040
Db 1981 GAAATCGTGGGACAAAGCATGATCNAACCGACGTTTAAAGCGGAAAAATTTCCATATT 2040
Oy 2041 caggcgagcgagcggtgatttcccgcaatgttgcgaagtggaagcgatgncattg 2100
Db 2041 CAGGCGGCGAGCGGTGATTTCGCGCAATGTTGCCAAAGTGGAAGCGGATTCNCATTTG 2100
Oy 2101 agcaatcgcgcccaagcagttttgtgtgcacacgcgcatcaaaagcatacaatctgata 2160
Db 2101 AGCATATCAGGCCAAGAGAGTTTGTGTGTCGACACCGCATCAAAAGCCATACATCTGACA 2160
Oy 2161 cgttcgagcctggaengtctgacaatgtgtgcgaanaaancatcccgacgataaagt 2220
Db 2161 CGTTGCGACTGGACNGGCTTGACAAATTTGTTCGAANMAANCAATTTACCGATTAAGTG 2220
Oy 2221 attgcttacttaagaacnagcagcagcagcagcagcagcagcagcagcagcagcagc 2280
Db 2221 ATTGCTTACTTAAAGCNCAGCTNAGCGGAGANTGTANAGCTNNCNCNTTMCAGTNTT 2280
Oy 2281 tnaaancnnggcnltgcmnaactnaangcaatcttaagtgcaaatgscgatacagct 2340
Db 2281 TNAANCTCNCNGGCGNTGCTNNACCTNANAGCAATCTTATGTCGAATGCGATACAGT 2340
Oy 2341 tatacagtcagccacaagcgcaccccaaaaacggaaccttagcctcgtggcgaatggccaa 2400
Db 2341 TATACAGTCAGCCACCAACGCCCAAAACGGCAACCTTAGCCTGTGGCAATGGCCAA 2400

Oy 2401 gnaacatttaalcgaagccacattaaacggaacncaatcggnlttcggcaatgtcttatt 2460
Db 2401 GCAACATTTAATCAACCCACATTTAAACGGCAACNATCGGNTTGGCATGTCTTCAATT 2460
Oy 2461 aatctaagcaacagcgcgcacaaaacgagcagcttgacgcttccgacaacgctaaagca 2520
Db 2461 AATCTAAGCAACAGCCCGGACAAAACGGCAGTCTGACGGCTTCCGAAACGCTAAGGCA 2520
Oy 2521 aacgttaagccatccgactcaacggaatgtctccctagccgataagcagatccat 2580
Db 2521 AACGTAAAGCATTCGCGACTCAACGGCAATGTCTCCCTAGCCGATTAAGGCGATTCAT 2580
Oy 2581 tttagaaacgcgcgttaccggaacacatcagcgcgacgaaganaacgatacacta 2640
Db 2581 TTTGAANAACGCCGCTTACCAGCAACTCAGCGGACAGCAAGANACAGCTTACACTTA 2640
Oy 2641 aaagacagcgaatggaacgctgcgttcagcagcgaattgaagcaattaaacctgacaac 2700
Db 2641 AAAGACAGCGAATGAGAGCTGCGCTCAGGCACGGAATTAAGCAATTTAAACCTTGACAA 2700
Oy 2701 gccacattacatcaattccgctatcgccagatctgtaggcgcgcgaacgcgagcag 2760
Db 2701 GCCACATTAACCTCAATTCGCTATCGCCAGATGCTGACGGCGCAAAACGGCAGN 2760
Oy 2761 gtgtcaagacgcgcgcgcgcgttcgcgcgttccctattatccgtataccgcgaact 2820
Db 2761 GTGTCAAGACAGCGCGCGCGCGCTTCGCGCGCTTCCGATTAATCCGTTACACCGCACT 2820
Oy 2821 tcggtagaatccgcttcaacaacgctgacggtlaaacggaataatgaaenltcaagaaca 2880
Db 2821 TCGGTAGAATCCCGTTCAACACGCTGACGTTAAACGGCAAAATTTGAAACGTAAGGAACA 2880
Oy 2881 ttccgcttattgtggaacitctggcctacggaagcgaacaaatgaactgvcggaagt 2940
Db 2881 TTTCCGTTTATGTGGAATCTTTGGCTACCGAAGCGCAAAATTTAAGCTGGCGGAAGT 2940
Oy 2941 tccgaaggaacttaccacttgcgtgcgtcaacaatacgcggaacgaacccglaagctgat 3000
Db 2941 TCCGAAGGNACTTACACTTGGCGGTCAACAAATACCGCAACGAACCCGTAAGCTCGAT 3000
Oy 3001 caattgaacgtatgtggaaggaagacaacaaacgcgtgcggaanaaccttaattcac 3060
Db 3001 CAATTGAACGTAATGGAAGGAAGAACAAACGCGTGTCCGAANAACCTTAAATTTACCC 3060
Oy 3061 ctgcaaaaacgaacagctgcgtgcgcgcgtgtgcgttaccactatccgcaaaagcgc 3120
Db 3061 CTGCAAAAACGACAGTGTGATGCCGCGGTGCGGTACCAACTCATCCGCAAAAGCGGC 3120
Oy 3121 gacttccgctgcaaatccggttcaagaagaagacttccgacaactcgcgcaagca 3180
Db 3121 GAGTTCCGCTGCATTAATCCGCTCAAAAGAACGCTTCCGCAAACTCCGCAAGGGA 3180
Oy 3181 gaagcacaacaaacgagcgaagaaagacaacgagcgaagccttgaacgcgtgtattg 3240
Db 3181 GAAGCACAACAAACAGCGGAAAGAACACAGCGCAAAAGCTTGCACGCGCTGATGGCGCT 3240
Oy 3241 ggcgcgcatgcccgcgaagaaagacagaagcgttgcgcgaacgcgcgcgcgcgcgcgcgc 3300
Db 3241 GGGCGCGATGCCGCGGAAGAACAGAAAGCGTTGCCGAACCGCGCGCGNGCGAGCGCG 3300
Oy 3301 gaaatgtcgcgcatatcagcagcgcggaaggaagaaacgcggttgaagcgcgataagac 3360
Db 3301 GAAATGTCTGGCATTAATGAGCGGAGGAAGAAACGGGTGACGGGGAATTAAGAC 3360
Oy 3361 agcgcnttggcgaacacagcgcggaaggaagaaacgcgcgcgcgcgcgcgcgcgcgcgcgc 3420
Db 3361 AGCGCNTTGGCGAAACAGCGGCAAGGGAACCGCGCGGNTACACACGCTTCCCGCGC 3420
Oy 3421 gcccgcnccgc 3480
Db 3421 GCCCGCNGCGCGCGCGGATTTGGCCGAACGCGACCGCCCAACCGCAACTCAACCCCA 3480

QY 361 aacgttgatttggtgcygaaggaagaaatccgcatacgcacccgtttcttaccaaatt 420
|||||
Db 94426 AAGGTGATTTTGGTGGGAAGGAAGAAATCCGATCGACACCGTTTTCTTACCAAAAT 94485
|||||
QY 421 gtgaagaaataaatttaagcctgcacaaatccacaccccttaacagcgatanaatg 480
|||||
Db 94486 GTGAAAAGAAATATATTAAGCCCTGACAAATTCACACCCCTTCAACAGCGGATTAACATATG 94545
|||||
QY 481 ccgcgtttgcataaattgtcacagatgcagaacccgtcgaatagcagatgcagtgagg 540
|||||
Db 94546 CCGCGTTTGATTAATATTTGTCACAGATGCAGAACCTGTGCAAAATGACGATGACATYAGG 94605
|||||
QY 541 ggggaatacctattccgataaagaaanaatalcccgagcgtgtccgcacatgcgtcagagac 600
|||||
Db 94606 GGGAAATACCATTTCCGATTAAGAAAATATCCCGAGCGGTGCCGATGGGTCAGAGAC 94665
|||||
QY 601 cactatgagctatgatgatgtgtaacaaacagcggaattatctactcctccgcgcagtgtta 660
|||||
Db 94666 CACTATTTGGGTTATGATGATGATGACAAACACGGCATTTATCTCTCCGCGCATGGTTA 94725
|||||
QY 661 attgcygcacatacacatacgcaggtgtggggaataatgcygtanttaattgaagcgc 720
|||||
Db 94726 ATTGGCGGCATATACATATATGCAAGGTTGGGGAATATATGCCGTAGTTTGAAGCGGC 94785
|||||
QY 721 gatgtgcacatgcgcacagcactatgacctatgcgcattgcaggtgcgcagagcagcagc 780
|||||
Db 94786 GATGTGCGCATGCTTCCGATTAAGCGCTATGGCCCTATGCGCATTCAGATGGCGGCGACAGC 94845
|||||
QY 781 ggttcgcacatgttttattatgcacaaacaaataatgagcgtgcacacagcgagttta 840
|||||
Db 94846 GGTTCGCCATGTTTATTTATGACAAACAAATATATGATGCTGCTCAACGGGAGTTTA 94905
|||||
QY 841 caaacccgctaaccttattccgcgcagggaaacaggttccagctgcatacgcagaaattg 900
|||||
Db 94906 CAACCGGCTACCCCTTATTCGGCAGGGAACGGTTTCCAGCGATACGGAATTTGG 94965
|||||
QY 901 ttctacgatacatttaacagagcgatacaccgtcctnttttgaacgcgcagtaac 960
|||||
Db 94966 TTCTAGATGACATTTACAGAGCGATACATACCTCTTTTGAACCGCGCAGTAC 95025
|||||
QY 961 ggcacatttctcttaccatacacaacacagctgcaggtacggtlaacagaaaccaaagaa 1020
|||||
Db 95026 GGACATTTTTCCTTATCATCCAAACAAACGATACGGGTACGGTACAGAAACCAACGAA 95085
|||||
QY 1021 aaggtntccaatcccaagcttaagtacagacagtcgcagctgttgaagaaatcttgaat 1080
|||||
Db 95086 AAGGTTTCCAAATCCAAAGCTTAAAGTACAGACAGTCCGACTGTTTGACGAATCTTTGAAT 95145
|||||
QY 1081 gaaactgataaagaacagatttacgcgcagaggggtgttaatacagtaacgtccaagttta 1140
|||||
Db 95146 GAAACTGATTAAGAACGATTTACGGGCGAAGGGGTGTTAATCATGATCCGTCAAAGGTTA 95205
|||||
QY 1141 aacaaagtgaaacaccttcttatacgtatcgcgaacgcgaactcatctatacaac 1200
|||||
Db 95206 AACCAAGGTGAAACCTTCTTTATCGATTAAGCGCAACGGCAACATCTTATCAAC 95265
|||||
QY 1201 aacatcaacaaagcgcggtgttgaatttgaagtgatttcaagcttcgcgtgaa 1260
|||||
Db 95266 AACATCAACCAAGCGCGCGGCGGCTGTGATTTGAAGGTGATTTTACGGCTCCGCTGAA 95325
|||||
QY 1261 aacaaagaaagtgcaagcgcgcggttcatatcagtgaaagacagtaacgttacttg 1320
|||||
Db 95326 AACCAAGAAAGGTGCGCAAGCGCGGCGGCTTCAATATAGTGAAGACGATCCGTTACTGG 95385
|||||
QY 1321 aaagttaacggtgtgcgaacgcgcgctgtccaaatgcgcaaggcagctgcacgtt 1380
|||||
Db 95386 AAAGTAAACGGCGTGGCAACGACCGCTGTCCAAATGCGCAAAAGCGACGCTGCAAGT 95445
|||||
QY 1381 caagccaaggggaaacaaagcgtcgatcagcgtgtgcagacggtacagcttaattggat 1440
|||||
Db 95446 CAAGCCAAAGGGGAACCAAGGCTGATCAAGCTGTGGCGGAGGTACATGATTTTGGAT 95505
|||||
QY 1441 cagcagcgacgcgtlaaagcgaaaaaacaagcctttagtgaaatcgctgttcagcgac 1500
|||||

Db 95506 CACAGGCGACGATTAAGGCAAAAAACAAGCCCTTTAGTGAATCGCTTGGTAGCGGC 95565
|||||
QY 1501 aggggtacggtgcgaactgaaatgcgaataatcagttcaaccccgacaactatcttcgac 1560
|||||
Db 95566 AGGGGTACGGTGCACCTGAATGCGGATATCAGTTCAACCCCGCAAACTTATTTGGC 95625
|||||
QY 1561 ttctgcgcgcagcgttcttgatttaacgggcatcgccttcgtttccacglatcaaat 1620
|||||
Db 95626 TTTCCGGGGGAGCGTTTGGATTTAAAGGGGCGCATTCGCTTGTCTCCACCGTATCAAAAT 95685
|||||
QY 1621 accgataagggcgatgtatgctatgcnatcgaatgccaacaaacatccacglttaccat 1680
|||||
Db 95686 ACCGATGAAGGGCGGATGATGTTCAATCATATATGCCCACCAACATCCACCGTTACAT 95745
|||||
QY 1681 acaaggaatgaaaglatlacacaacccagtgtaagaatalcaatagacttaacagc 1740
|||||
Db 95746 ACAGGAATGAAGATTAATACACACCGAGTGTAAAGATATCATATGACTTAATTTACAGC 95805
|||||
QY 1741 aaagaatgtcctacacaggttggtttgtagagaagaatagacccaacgaacggcg 1800
|||||
Db 95806 AAAGAAATGTCTTACACGCTTGGTTTGGGAGAAAGATACGACCAAAACGAACGGGCGG 95865
|||||
QY 1801 ctcaacctgtttacacagccgcgcagagaagacgcaccngctgtcttccgcgcgagaca 1860
|||||
Db 95866 CTCACCTTGTTTTACAGCGCCGCCGAGAAAGACCGCACCCCTGCTTCCGCGGAGACA 95925
|||||
QY 1861 aatttaaacgcgaacatcacgcgaacaaacgcgaacgttlttcaagcgcgacgcgaca 1920
|||||
Db 95926 AATTTAAACGGGCAACATCACGCAAAACCAACCGCAAACTGTTTTCAGCGGCAAGCGACA 95985
|||||
QY 1921 ccgcagccctacaatcaattagaagcgggtgtgtcaaaaatggaagatccccaagga 1980
|||||
Db 95986 CCGCACGCTTACAAATCATTTTAAAGAACGGGTGTCTCAAAATGGAAGTATCCACAAAGGA 96045
|||||
QY 1981 gaaatcgtgtgggaacacgcactgtaacccgcgcagcttlaaagcgcggaatltccatatt 2040
|||||
Db 96046 GAAATCGTGTGGAGCAACGACATGATCAACCGCACGTTTAAAGCGGAATTTCCATAT 96105
|||||
QY 2041 cagggcgggcgaggggtgatttccgcgaatgtgtccaaaggtggaagcgatgnaattg 2100
|||||
Db 96106 CAGGGCGGGAGGGCGGTGATTTTCCGCAATGTTGCCAAAGTGAAGGCGATTTGGCATTTGG 96165
|||||
QY 2101 agcaatcacgccccaagcagtttgtgtgcgacgcgcatacgaacatacatctgtaca 2160
|||||
Db 96166 AGCAATCACGGCCCAAGCATTTTGTGTGCGCACCGCATCAAAAGCATATCTGTATA 96225
|||||
QY 2161 cgttcgcagctggacngtctgcacaaatcgtgtcganaaaanacatlaacgcagataaagt 2220
|||||
Db 96226 GCTTCGAGACTGGACCGGCTGTGCAAAATGTGTGCAAAAAACCATTAACGACGATAAAGTG 96285
|||||
QY 2221 attgcttcatgactaaagacnagcgnlnagcggcatalgtnagctlnmcnatacgnlmt 2280
|||||
Db 96286 ATTGCTTATTTGACTATAGACCGCATCTCAGCGCAATGTCAAGCTTGGCGATCACGCTCAT 96345
|||||
QY 2281 tnaaantcnngggcmtlgnmcactlnaangcgaatctttagtgcgaatlgcgatacaagt 2340
|||||
Db 96346 TTAATCTCACAGGGCTTGCCACACTCAACGGCAATCTTATAGTCAAAATGGCATACGCT 96405
|||||
QY 2341 tatcagtcagccaacaaagccaacccaaagggcaaccttagcctcgtgtgggaatgcgca 2400
|||||
Db 96406 TATACAGTCACGCCAACGACCCCAAAACGGCAACCTTAGCTTCGTGGGCAATGCCCA 96465
|||||
QY 2401 gcaacatttaacagccacatlaaagcggaacnatacgnlttcgggaatgcttcatatt 2460
|||||
Db 96466 GCAACATTTTATTAAGCACATTTAAACGCGACACACATATGCTTCCGGCAATGCTTCATTT 96525
|||||
QY 2461 aatctaaagcaacacgcgcgcaacaaacgcgcagcttgacgcttccgacaacgcctaagca 2520
|||||
Db 96526 AATCTAAGCAAAACGCGCGCACAAACGCGAGTGTGACGCTTTCGACACAGGCTTAAGGCA 96585
|||||
QY 2521 aacgttaagcatttcgcacatacgcgcaatgtcttcctatgcgataagcagtatcat 2580
|||||

Dh 96586 AACGTAAAGCCATTCCGACATCAACGGCAATGTCTCCCTAGCCGTAAGCAGTATTCAT 96645
Qy 2581 ttgtgaaacagccgctttaccggaacattcagcgcaagaaanacagattacattta 2640
Dh 96646 ttgtgaaacagccgctttaccggaacattcagcgcaagaaanacagattacattta 96705
Qy 2641 aaagacaggaatgagcgctgctcgaagcaagaaattagcaatttaaacctggaacac 2700
Dh 96706 AAAGACAGGAATGAGCGCTGCGCTCAGCAGCAAGGAATTAAGCAATTAACTTGACAC 96765
Qy 2701 gccacattacattcaattccgctcctacgcacagatgctgcaagcgcaaacccgcaag 2760
Dh 96766 GCCACATTACACTCAATTCGCCCTATCCGCCAGATCTCAGCGCGCAAAACCGCGCAGT 96825
Qy 2761 ggttcagacacgcgccgctgctgctgctgctgctgctgctgctgctgctgctgct 2820
Dh 96826 gttctcagacacgcgccgctgctgctgctgctgctgctgctgctgctgctgctgct 96885
Qy 2821 tcggtagaaatccgctttcaacacgctgacgtaaacgcaaatlgaaacngtcaagaaaca 2880
Dh 96886 TCGGTAGAAATCCCGCTTTCACACGCTCAGCGTAACGCGAAATTGACAGCTCAAGAAC 96945
Qy 2881 ttcgccttattgctcgaacattcctcgcgtacgaagcgcaaacattgaagctggcggaagt 2940
Dh 96946 TTCGCTTTATGTGCGAACTCTTCGCTACGAGCGACAAATTTGAACCTGGCGAAAGT 97005
Qy 2941 tcggaagaaacttaacacctggcggttaacaatacgcgcaacgaaacccgtaagctgcat 3000
Dh 97006 TCGGAAGGCACTTACACCTTGGCGGTCAACATACGCGCAAGAACCCGTAACCTTCAT 97065
Qy 3001 caattacggttagtggaaggaagaaacaacaacacgctgtccgaaaccttaatttccac 3060
Dh 97066 CAATTACGCTAGTGGAAAGGAAAGAACAAACCGCTGTCCGAAACCTTAATTTCAC 97125
Qy 3061 ctgcaaacagacacgctcgaatgctcgcgctgctgctgctcgaactcgcgaagaagcgc 3120
Dh 97126 CTCGAAACAGCAACGTCGATGCGCGCGCTGCTTACCACTCATCTCCGCAAAACGCGC 97185
Qy 3121 gaattccgctgataatccggtcaagaagaagaagcttcgaaacacctcgggaagca 3180
Dh 97186 GAATTCGCTGCTAATCCGCTCAAGAACAGAGCTTTCGACAACTCCGCAAGCA 97245
Qy 3181 gaagccaaacacagcgcggaacacacgcaacgcttgaacgctgctgctgctgctgct 3240
Dh 97246 GAAGCCAAACAGCGGCGGAAAGAACACACGCGCAACCTTGACGCGCTGTCGCGC 97305
Qy 3241 gggcgcgatgctcgcgcgaagaagcgttgcggaacccgcgccggaagcgagcg 3300
Dh 97306 GGGCGGATGCGCGCGCAAAAGACAGAAAGCGTTGCCGAACCGCGCGCAGCGAGCGG 97365
Qy 3301 gaaatgctgcatattgctcgcggtggaagaagaagaagaaggtgctcagcgagataagac 3360
Dh 97366 GAAATGCTCGCATTTATGCAAGCGGAGAAAGAAAGAGGCTGACGCGATTAAGAC 97425
Qy 3361 agcgcttggcgaaacagcggaagcggaacccgcgctgacacgcttcccccgcg 3420
Dh 97426 AGCGCCTTGGGAAACAGCGGGAAGCGGAAACCGCGCGGTACACACGCTTCCCGCGC 97485
Qy 3421 gcccgmgcgccgctggaatttgcgcgaacgcgaccccaacgcgcaacctaacccca 3480
Dh 97486 GCCCGCGCGCGCGCGGATTTGCCGACCGCGACCGCCCAACCTCAACCGCCCA 97545
Qy 3481 ccggaagcgacacgcatgacgcttattgccaatagcgtttgaagaaatttccgcaag 3540
Dh 97546 CCGGACGCGCACTGATACGCGCTTATGCCATAGCGGTTTAGTGAATTTCCGCGCAG 97605
Qy 3541 ctcaacagcgcttctcgcctgacaggaagaaatggaacgctgctgctgcaagacgcgcg 3600
Dh 97606 CTCACAGCGCTTTCGCGCTACAGGACGAATTTGACCGCTGTTGCCGAACCGCGCGC 97665
Qy 3601 aagcngtttggacaagngatccggaacccaacactacgcttggaaatctccgcg 3660
Dh 97666 AACCGCTTTTGACAAAGCGGATCCGGGACCAACCAACTACCGTTGCAAGATTTCCCGC 97725

Qy 3661 gctaacgcgaacaaacgcgacctgctgcgaatcgtatgcaagaaaaactcgcgcgagcg 3720
Dh 97726 GCTACCGGCAACAAACGACCGCTGGCGCAATGATGACAGAAAACTTGGACGCGG 97785
Qy 3721 cgcctcgcgcatcctgttctcgcgaacacgcggaacccgaacncttcgacgcgcgcatcgcg 3780
Dh 97786 CCGCTCGGATCCCTGTTTTCGCAACACCGGACCGAAACCTTGCAGACGCGATCGGC 97845
Qy 3781 aactcggaacgcttgcccaacgcgcgcttctggggaatacagcatcgcgaagttcgac 3840
Dh 97846 AACTCGGACGCGCTTCCCAACGCGCGCTTTCGGGCAATACGCAATCGCGAGTTGAC 97905
Qy 3841 atcgcatagcaagcgcgcggttttaagcagcgcaantcintcgaacgcgcatcgcgagc 3900
Dh 97906 ATCGGATCAGCACGCGCGCGGTTTACGACGCGGAGTCTTTCAGACGCGATCGAGCG 97965
Qy 3901 aaaaatccgcgcgcgctgctgcatcgaatcgaagcaacgatacgcgcgcggttctgcgcg 3960
Dh 97966 AAAATCCGCGCGCGCTGCTGCTACGATTCAGGCAATCAGCAGATACCGCGGTTTCGCGC 98025
Qy 3961 ggaattcgatcgaacgcgtaacatcgcgcaacgcgcgctatttcgtccaaaagcgatcac 4020
Dh 98026 GGAATTGCAATCGAACCGTACATCGCGCAACGCGCTATTTCGTCCAAAAACGCGATTAC 98085
Qy 4021 cgtacgaaacgctcaaatcgcgaaccccgctctgctgcttcaacacgntacgngcgagc 4080
Dh 98086 CGCTAGAAACGTCATATCGCACACCCCGGTTTCGTTCAACCGCTACGCGCGCGC 98145
Qy 4081 ataaagcagatattcaatcaaacgcgcaacacacatnccaacncttattagc 4140
Dh 98146 ATTAAGGCAAGATTATTCATTCAAACCGCGCAACCATTTCCATTCACGCTTATTGAGC 98205
Qy 4141 ctgtccataacgagatgcgcttgcggaagaagctccgaacacgcgctcaatccgngtattg 4200
Dh 98206 CTGTCTATACCGATGCGCTTTCGCGCAAGCTCCGAACGCGCTCAATACCGCGCTATTG 98265
Qy 4201 gctcagatctcgcgaacacccgcagctgcggaatgagcggtgagcgcaaacgcaaatc 4260
Dh 98266 GCTCAGATTTTCGCAAAACCGCGCAGTGGCAATGGGCGCTGAACCGCGCAATCAAGT 98325
Qy 4261 ttcaagctgtcctcgcgcgctgctcgcgcgaagaagncgcgaacgaggaagcgcaacacgcg 4320
Dh 98326 TTCACGCTGTCCCTCAACGCTGCGCGCAAGGCGCGCAACTGGAAGCGCACACGAC 98385
Qy 4321 gcgcgcatcaaatagctaccgctgttaa 4350
Dh 98386 GCGGCAATCAATTAGCTACCGCTGTAA 98415

RESULT 3
NME296276 4350 bp DNA linear BCT 14-JAN-2002
LOCUS NME296276
DEFINITION Neisseria meningitidis app gene for adhesion penetration protein, strain 23515.
ACCESSION AJ296276
VERSION AJ296276.1 GI:11071862
KEYWORDS adhesion penetration protein; App gene.
SOURCE Neisseria meningitidis
ORGANISM Neisseria meningitidis; beta subdivision; Neisseriaceae; Neisseria.
REFERENCE 1 (sites)
AUTHORS Hadi,H.A., Wooldridge,K.G., Robinson,K. and Ala'Aldeen,D.A.
TITLE Identification and characterization of App: an immunogenic autotransporter protein of Neisseria meningitidis
JOURNAL Mol. Microbiol. 41 (3), 611-623 (2001)
MEDLINE 21424673
REFERENCE 2 (bases 1 to 4350)
AUTHORS Ala-Aldeen,D.A.
TITLE Direct Submission
JOURNAL Submitted (27-OCT-2000) Ala-Aldeen D.A., Microbiology, University of Nottingham, Molecular Bacteriology and Immunology Group,

FEATURES		University Hospital, Nottingham NG7 2UH, UNITED KINGDOM	
source	Location/Qualifiers		
	1..4350		
	/organism="Neisseria meningitidis"		
	/strain="23515"		
	/db_xref="taxon:487"		
	/country="Saudi Arabia"		
	/note="serogroup A"		
gene	1..4350		
	/gene="app"		
	1..4350		
CDS	/gene="app"		
	/codon_start=1		
	/transl_table=11		
	/product="adhesion penetration protein"		
	/protein_id="CAC14670.1"		
	/db_xref="GI:11071863"		
	/db_xref="SPTREMBL:O9F325"		
	/translation="MKTIDKRTTETHRKAPKTRIRFSPAYLAIICLSFGLIPQAMAGH TFEGLNYOYRDFPAENKKEFAVGAKDIEYNNKGEIVYKSMTPAKMIDFSEVRNGVA ALVGPQYIVSVANGNGYNNVDEGAEGRNPDQHFYSQIYKRNMYKPDNSHPYNDYIM PLRHKFVTDIAEPEVMTSDMRGNTYSDKEKPEPERVIRIGSGHHYRYYDDKRGDLSYSGA WLIIGNTHMGGNGNGVSLSGDVIRHANNYQPMPTAGAGDSSGPMPTIDKTNKKLL NGVLOTGYPSGRNGFOLIRKDMRYDDIYRCDTHRVFPEPRNSGHPSEYNNNGTGT VETHEKYSNPKLEVOYVAFLDESLNETKEEYVAGGVNQVYRPTLNKNGVANDRL SKIGTGLHVOAKGNGOSISVGDFTVILDOQADDGKQKQAEISGLVSRGVQJOLA DNOFNPDKLIEFGREGRLDLNGHLSFHRILQMTDEGAMVNNHATITTYITIGNESI TOPSGKNINRLYKSKIEAYNGWFEKDTKTNGRLVYVOPAEEDTLILSGGTINIG NITQNGKLFESGRPTPHAYNHLGSGMSMEGIPOGETIYMDNDMIRKTRAEFPHIOG GOAVSRVAVYEGDMHLSNHOAVPGVAPROSHITCTSDMTGLTNCVKTITDYK IASLTKTDISGNSLADHAHLNLTGLATLNGMLISANGDRTYVSHATONGNLSLVN AQTFTNQATVLTGNTISGNSAFNLNNNAONGSLTLDNAKANVSHSALNGVSLADK AOFHEHNRFTGQLSGSKDFTALHKDEWTLPSGELNLIDNATITINSAYRIHDA GAQGTGSVDTPRRRSRSLSVTPPTYSRENTLTLNKLNGOCGFRFMSFEGYRS DKLIAESSEGTITLVANNNGNEPYSLOLTYVEGDKNPLSNNLFTLONEHVDGA WRYOLIRKDGEPRLHNPVKEOELSDKICAEAKKQAEKNNAOSLDLIAAGRAAKET ESVAEPARAGGNGYIMQAEERKRVQADKDSALAKQREPRFTVAPRRARAV DLPOPOPPOPORDLISRTANSGLSEFSATLNAVFAVEDLRVAEDRRNARV TSLRDTKHYSODEFRAYRQQTDLRQIGQKNLIGSGRVILFESHNTENTFEDIGINS ARLAHGAFFGGYIGIRPDIGISTGAFSGSLSDIGIGLIRRVLYLGIQARAFEG GFGIEPYIGATRIYEVOKADRYENYENIATPGLAFNRYNGITADYSFKRAOHSTIPY LSLSTYDASGKRVTRVMTVAIADQDFKTRSAEMGVNNAELKFTLSLHAAAKGPDL AQSAGITGLGRW"		
BASE COUNT	1236 a 1185 c 1073 g 856 t		
ORIGIN			
Query Match	98.5%:	Score 4284.8:	DB 1; Length 4350;
Best Local Similarity	98.5%:	Pred. No. 0:	
Matches 4286;	Conservative 0;	Mismatches 64;	Indels 0; Gaps 0;

1	atgaaacaacgcgacaacgcgacaacgcgacaacgcgacaacgcgacaacgcgacgc	60
2		
3		
4	1 ATGAATAAACACGCAAAACGGAACAACCGAAACACACCAGGCGGCGCGC	60
5		
6		
7	61 atccgcttcgcgcgtactactagcatatgctctgttgctgcgacattctcccaagct	120
8		
9	61 ATCCGCTTCTGCGCTCGCTACTTATGACCATATGCGCTGTGTCGCAATCTTCCCAAGCT	120
10		
11	121 tggcgaggacacactatttcgcgatccaactcaactactactcgcgcgacttgcgcgaaat	180
12		
13	121 TGGCGGGGACACACTATTTCGCGCATCACTAACATATGATGCGGACTTGGCCGAAAT	180
14		
15	181 aaaggcaagtctgcgaatctcgggcgaaagatactgaagttatacaaaaaaagggaagt	240
16		
17	181 AAAGGCAAGTTTGCACTGCGGGCGAAAGATATTTAGGTTTACAACAAAAAGGGGAGTTG	240
18		
19	241 gtccgcaaatcaatgaacaaaggcccgatgattgattcttcctggtggtcgcgtaagcgc	300
20		
21	241 GTCCGCAAAATCAATGACAAAAGCCCGCATGATTATTTTCTGTGGTGTGCGCGTAAGGCG	300
22		
23	301 gtggcgcatctgtgtggcgatcaatatattgtgaacgttgcacataacgcgcgtataac	360

Db	301	GTGGCGGCAATTGGTGGCCATCAATATTTGTGAGGTGGCAATTAACGGCGCTATAAC	360
QY	361	aacgttgatcttgytgcggaagaagaaalcccgatcagaccgttcttctcaaat	420
Db	361	AACGTGATTTGTGTCGGAGAGAAATCCGATCAGACCGTTTCTTACCAATT	420
QY	421	gtgaaagaataatataagcttgcaatcaacccttacaagcgcatancatalg	480
Db	421	GTGAAAGAAATATTAAGCTTGCAATTTCAACCCCTTACACGCGCATTTACATATG	480
QY	481	ccgcgtttgcataaatcttgcacagatgcagaaacctgcgaatgcagatgcagag	540
Db	481	CCGCGTTTGCAATAATTTCTTCACAGATGCAAGACCTGTGAAATGACGACTGCATGAGG	540
QY	541	gggaatcactatccgataaagaanaatacccgagcgigtccgcatcgcgtcaggaac	600
Db	541	GGAATACCTATTCCGATTAAGAAAATATCCCGAGCGTGTCCGATCGGCTCAGGACAC	600
QY	601	cactatlgcgftatgatgatgatacaaacgcgcgcatattactctaccgcgcgcatggtta	660
Db	601	CACATTTGGCGTTATGATGATGATCAAAACGCGGATTTATCCTACTCGCGCATG6TTA	660
QY	661	attgcggaatacacatatacagaggtltgaggaaataatgagcgtanltagttgagcgcg	720
Db	661	ATTGCGGCAATACACATATGCAAGGTTGGGAAATTAATGGGCTATTGTTGACCGCG	720
QY	721	gatgcgccatgccaaagcactatgcgcctatgcgcgatgtcagglgvcgagcgagcagc	780
Db	721	GATGCGCCCATGCCAACACATATGCGCCCTATGCGGATTTGACGATGCGGACGCGACACG	780
QY	781	ggtcgcaaatggttattatctacacaaacaaataaagtgcgtctcaacggaagtta	840
Db	781	GGTTCGCCAATGTTATTATTATACAAACAAACATTAATGCTGCTCAACGGAAGTTTAA	840
QY	841	caaacgcgtacccttattccgcgcaagggaanaacggttccagctatatacgaanaagt	900
Db	841	CAAAACCGGCTACCTTATTTCCGGGCAAGGAAACGTTTCCACGCTGATACGAAAGTATGG	900
QY	901	ttctacagatgacatttacaagagcgatatacaataccgtctcttcttgaacggcgagtaac	960
Db	901	TTCTATGATGACATTTACAGAGCCGATACATACATCCGCTCTTTTGAACCGCGCACTAAC	960
QY	961	ggaacatttctccttaccatccacacacaaacggttacgaggttacggttaacagaaacacgaa	1020
Db	961	GGACATTTTCTCTTATCATCCAAACAAACAGGTACGGGTAAACGAACAAACGAA	1020
QY	1021	aaggtttcccaatccaaggttaaggtacagacagtcgactgtttgaacgaatcttgaat	1080
Db	1021	AAGGTTTCCCAATCCAAACACTTGAAGTACAGACAGTCCGACTGTTGACGAATCTTTGAAT	1080
QY	1081	gaaactgataaagaacaggtttacgcgcgacaggggtgttaatcagttacgttccaagtta	1140
Db	1081	GAAACTGATAAAGAACAGTTTACCGCGCAGGGGTGTTAATCACTATCCGTTCAAGTTAA	1140
QY	1141	aacaaagtgaaacacttcttctatcgatlcgacgcaacgycgaactcatctataaac	1200
Db	1141	AACAACGCTGAACACTTTCTTTATCGATTACGGCAACGGGAAMCTCATCTTATCAAAAC	1200
QY	1201	aactcaacaaagcgcgaggggtttgtatcttgaaggtgattttaaggtctgcctgaa	1260
Db	1201	AACATTCACCAAGCGCGGGGGGTTGTATTTTGAAGGTGATTTTACGGTCTCGGCTGAA	1260
QY	1261	aacaacgaaacgttgcaagcgcgagggcggttaatacagtgaaagcaagtaacgttaactgg	1320
Db	1261	AACAACGAACAGTGGCAAGCGCGGGGCTTATATTCAGTGAAGACAGTACCGTTACTGG	1320
QY	1321	aaagtaaacggttgcgaacgacgcctgtccaaatctcgcgaagcgaacgctgcagct	1380
Db	1321	AAAGTAAACGGCGTGGCAAAAGACCGCTGTCCAAATCGGCACAAAGGACGCTGCAGCTT	1380
QY	1381	caagccaaggggaaacacgaagctcgatacaggttggcgacggttacagttacttggat	1440

Db 1381 CAAGCCAAAGGGGAAAAACAAGGCTCGATCAGCTGGGCGACGCTACGATCTTTGGAT 1440
Qy 1441 cagcagcgagcgatataaaggcaaaaacagcccttagtgaaatcggtgttntcagcgcc 1500
Db 1441 CAGCAGGCAACACATTAAGCAAAAAACAAGCCTTTAGTGAATCGGTGTGTCGCGGC 1500
Qy 1501 aggggtacggtgcaactgaatgcgcatatcaatcccccgaacactcatttcggc 1560
Db 1501 AGGGGTACGGTGCAGTCAATGCGCATATCATGTTCAACCCCGCAAACTCTATTTCGGC 1560
Qy 1561 ttccgcgcgagcgatttgaatttaacgaggatcgcgttcgttcacggtatccaat 1620
Db 1561 TTTCCGGCGCGAGGTTTGATTTTAACGGGCACTTCGTTGGTTCACCGATTCAAAAT 1620
Qy 1621 accgataagaggcgatgattgtnacataatgcacaaacaatccaacgttaccat 1680
Db 1621 ACCGATCAAGGGGCGATGATTTGTCATCATATATCCACACAAACATCCGTTACCAT 1680
Qy 1681 acagggaatgaagtattacacacgagtggtgaagaaatacaatagaacttaatacagc 1740
Db 1681 ACAGGCAATGAAGTATTACCAACCGATGTTAAGATATCATATGACTTAAATTACAC 1740
Qy 1741 aaagaaattgctactacaacggtltggtltggtcgagaagaatacagccaaaaacgaa 1800
Db 1741 AAAGAAATTGCTTACACAGGTTGGTTGGCGAAGATACGACCAAAACGAGGCGG 1800
Qy 1801 ctcaacctgttaccagcccgcccgagaaagacgcaaccngctgtcttcgcggaaga 1860
Db 1801 CTCACACTTGTTTACGAGCCCGCGAGAAAGACCGACCTCTCTTCGCGGGAACA 1860
Qy 1861 aatttaacggcaacatcacgcaacaaacgcaaacgtlttccaagcgcgagcgaca 1920
Db 1861 AATTTAACGGCAACATCACGCCAAACAAACGCAAACTGTTTTCAGCGGACAGCGACA 1920
Qy 1921 ccgacgagctactaatcttaggaagcggtggtgcataaataatgaagtatacccaagaa 1980
Db 1921 CCGACACCTTACATATTTAGGAAGCGGTTGTCMAAAATGGAAGGTATCCACAAAGA 1980
Qy 1981 gaaatcgttgggacaacgactgatatcaccgacgtttaagcggaataattccatct 2040
Db 1981 GAAATCGTGTGGACACGACTGATGTACACCGCATTTTAAACGGCAAAATTTCCATAT 2040
Qy 2041 caggcgagcgagcggtgatttccgcgaatgtltgcgaagtgtgaagcgatgtgncattg 2100
Db 2041 CAGGCGGCGAGCGGCGTATTTCCGCAATGTTCGCAAAAGTGGAAGGCGATTTGGCAT 2100
Qy 2101 agcaatacagcccaagcagttttgtgtgtcgacccgcatcaagcatacatctgtaca 2160
Db 2101 AGCAATACAGCCCAAGCAGTTTGTGTGTGCGACCGCATCAAAAGCCATCATCTGTACA 2160
Qy 2161 cgttcgagctgagcnggtctgacaaatgtgtcgaaaaaancattacgacgataaagt 2220
Db 2161 CGTTCGACCTGGACGGGTCTGTCAAAATGTGTGCAAAAAACCATTTCCACGATTAAGTG 2220
Qy 2221 atgtctcatctgactaagaacngacnlnagcggcantgtlnagntnmcnathacnlnnt 2280
Db 2221 ATTGCTCATTTAGCTAAGACCGACGATCATGAGCGGCAATGTACGCTTCCGATCAGCTCAT 2280
Qy 2281 tnaaancncnnggngntgcmnactnaangcaatcttagcagaatgtgcatacgt 2340
Db 2281 TTTAATCTCAAGGGGCTTGCACACTCAAGGCAANTCTTAGTCAAAATATGGCGATACAGT 2340
Qy 2341 tatacagtcagcacaacgcccacaaacggaacacttagccctgttggtgcaatgtgccaa 2400
Db 2341 TATACAGTCAGCCACAAAGCCACCAAAAGGCAACCTTAGCTGTGGCATATGGCCAA 2400
Qy 2401 gcaacatttaatacgaacacatlaaagcgcaacncaatcggnltcggtgcaatgtctta 2460
Db 2401 GCAACATTTAATCAAGCCACATTAAGACGCAACATCGGCTTCGGGCAATGCTTCAAT 2460
Qy 2461 aatcagaacaaacgcccagacaaacggtcattcgaacgttcggaacgtaagaa 2520
Db 2461 AATCTAAGCAACAGCGCGACAAAGCGCATGTGACGCTTTCGACAAACGCTTAAGCA 2520

Qy 2521 aacgttaagccattccgaactcaacggaatgtctccctagccgaataagcagta ttccat 2580
Db 2521 AACGTAAAGCCATTCCGACATCAACGGAATGTCTCCCTAGCCGATTAAGGCGATTTCCAT 2580
Qy 2581 ttgnaaaacagccgcttaccggaacaactcagcggaacgaaganaacagatcactta 2640
Db 2581 TTTGAAACAGCCGCTTTACCGGACAACTCAGCGGACGAAGGATATACGATTAACACTTA 2640
Qy 2641 aaagaacggaatggaagcgtccggtcaagcaggaatlaagcaattaaacttgaacac 2700
Db 2641 AAAGACAGCAATGAGAGCTGCGCTCAGGACGGAATTAAGCAATTAACCTTGACAAAC 2700
Qy 2701 gccacattacatcactcctccgctcatalcgcaagatctgcaagcgcgcaaacgycagn 2760
Db 2701 GCCACATTTACTCAATTCGCGCTATTCGACAGATGCTGCAGCGCGCAAAACGGCGAGT 2760
Qy 2761 gttcagacacgcccgcgcttcaacagcttagcgtlaaagggaacttgaacnftcaagaa 2820
Db 2761 GTGTACAGACAGCGCGCGCGCTTCGCGCGTCCCTATTATCCGTTACACCGCAACT 2820
Qy 2821 tccgttactgtcggaactcttcggtacccgaagcgacaacttgaagcttgcggaaga 2880
Db 2821 TCGGTAGAATCCGTTTCAACACGCTGACGATAACGCAATTAAGCGTCAAGAAACA 2880
Qy 2881 ttccgcttactgtcggaactcttcggtacccgaagcgacaacttgaagcttgcggaaga 2940
Db 2881 TTTCCGCTTTATGTGGAACCTTTGCGGTACCGAAGCGCAAAATTAAGCTGGCGGAAAGT 2940
Qy 2941 tccgaaggaacttaacacttgcggtgcacaaatccgggaacgaacccgtgaagctgat 3000
Db 2941 TCCGAAGGCACTTACCTTGGCGGTACACATACCGGCAACGAACCGTAAAGCTTGAT 3000
Qy 3001 caattgaacgtagtgggaaggaagacacaaacccggtccggaacacttaattccacc 3060
Db 3001 CAATTGACGTTAGTGGAAAGGAAGACACAAACCGCTGTCCGAATCTTAATTTCAAC 3060
Qy 3061 ctgcacaaacgaacacgtcgaatgcggtgcgttgcgttaccacactcaatccgaaga 3120
Db 3061 CTGCAAAACGAACACGTCGATGCCGCGGTGCTTACCAATCTATCCGCAAAACGCGC 3120
Qy 3121 gaattccgcttgcataatccggtcaagaagaagcgttccggaacacttcggaagga 3180
Db 3121 GAGTTCCGCTGATATTCGCTCAAAACGAAGAGCTTTCCGAATAACTGGCAAGGCA 3180
Qy 3181 gaagcacaacaaacagcggaacaaacgaacgcaacgcttgaacgctgaatttcggtcc 3240
Db 3181 GAAGCCAAAAACAGCGCGGAAAAACAAACCGCAAAACCTTGAGAGCGGTGATTTGGCGC 3240
Qy 3241 gggcgagatgcgcgcgaacaaagcgaacggtltgcgaacggccgngcgagcggtg 3300
Db 3241 GGGCGCATGCGCGCGGAAAGACAGAAAGCGTTGCCGAACCGCCCGGACGAGCGGG 3300
Qy 3301 gaaatgtcgcattatgcgaagcggaagaaagaaacgggtgtcagcggaataaagac 3360
Db 3301 GAAATGTGCGCATTAATGACGAGGAGGAAGGAAGAAACGGGTGACGCGGATTAAGAC 3360
Qy 3361 agcgnttgcggaacacgagcgaacggcggaacccggtatcacacgcttcccccgc 3420
Db 3361 AGCGCTTGGGGAACACGCGGAACCGGGAACCGGCGCGGTACCAACCGCTTCCCGCG 3420
Qy 3421 gcccgngcgccgcgggatttgcgcaacccgcaaccccaacggcaactcaaccgca 3480
Db 3421 GCCCGCCCGCGCGCGGATTTGCGCAACCGGACCGCCCAACCGCAACCTCAACCCCA 3480
Qy 3481 ccgcaagcgagccgtatgaatgcgttatgcgaatagcgtttagtgaatcttccgcaag 3540
Db 3481 CCGCAGCGGACCTGATACGCGGTATGCAATACGCGTTTAGTAATTTTCGCGACAG 3540
Qy 3541 ctcaacagcgttcttcggtacaggaacgaatlgacccggtgtltgcggaagacgcgcg 3600
Db 3541 CTCAACAGCGTTTTCGCGGTACAGACGAAATTGACAGCGGTGTTCGCAAGACCGCGC 3600

BAPF COUNT	1266	a	1163	c	1066	g	879	t
ORIGIN								
Query Match	80.6%; Score 3508; DB 6; Length 4374;							
Best Local Similarity	88.4%; Pred. No. 0;							
Matches 3876; Conservative	0; Mismatches 462; Indels 48; Gaps							
OY	1	atbaaaacaacgcgaacgaacagcagaccggaacacacccgnaaagccccgaaaaccgycgc	60					
Dd	1	ATGAAACCAACCACCAACAACGCAGCACCCGAACACACCGCAAGGCCCGCAAAACCGGCCGC	60					
OY	61	atccgccttcgcgcctactacttagcatatgccttgcgttcggacattttcccgaagt	120					
Dd	61	ATCCGCTTCGCGCCTACTACTTAGCATATGCTGTCTGTTGGCATATTCTTCCCAGAAGC	120					
OY	121	tgaggcgggacaacattattcggcalcaaacatacctaactatcgcaccttgcgaaat	180					
Dd	121	TGGCGCGGACACACTTATTTCGGCATCACTACCATTAATACTATCGGCACTTCCGCAAAAT	180					
OY	181	aaaggcaaatcttcgacgtcggcggaagaatattgagtttaacaaaaaagggaattg	240					
Dd	181	AAAGGCAAGTTTGCACTCGGGGGAAGATTTGAGTTTAAACAAAAGGAGGAGTGG	240					
OY	241	gtcgcgaacalcaatgacaaagcccgcgaatgatatttttcgttgtcgtqtaagc	300					
Dd	241	GTCGCAATTCATGACAAAGCCCCGATGATTTGATTTTTCTGTGTGCGGTAAACGC	300					
OY	301	gtggcgcaattggttgvggcgaacaatatattgagcgttgacaaacagcggcataac	360					
Dd	301	GTGGCGCATTTGGTGGCGCATCAATATATTGACGCTGGCAATACGCGCGCTAATAC	360					
OY	361	aecgttgactttgltgcggaaggaanaatcccgatccagcacgcttttctaccaaat	420					
Dd	361	AACGTTGATTTTGTGTCGCGGAAGSMAAATCCCGATCAACTCGTTTACTTAAAAAT	420					
OY	421	gtggaagaataatataagccctgacaatta-----caacctacaacgycgatanat	477					
Dd	421	GTCGAACGGAATTAATTATTAAGCAGGACTAAGGCCATCTCTTATGGCGGATTAATCAT	480					
OY	478	atgcgcgctttgcataaatttgtaacagatgagaacccgttcogaaatgacgagtactg	537					
Dd	481	ATGCCCGCTTTCATTAATTTGTACAGATGAGAACCTGTGAATATACCAAGTTATATTTG	540					
OY	538	agggggaataccatattccgatalaagaaaataatcccgagcgtgtccgcgatcgcgaqa	597					
Dd	541	GATGGCGGAATTAATATGATCAAAATTAATTAACCTGACCGGTGTGATTTGGGCGAAGC	600					
OY	598	caaccatattgacgttatagatga-----caaacgcgcgacttatcctaactcgcgc	651					
Dd	601	AGGCATATTTGGCATGTGATGAATGATGACCCCAATTAACCGCAAAAGTTCAATATCATATT	660					
OY	652	gca-----ttgtaattbgcggaataacatatgcagcgtttgggaataat	699					
Dd	661	GCAAGTGGGTATTCTTGGCTCGTTGTGTCGCAATACCTTTGCACAAAATGATACAGTGTGT	720					
OY	700	ggcgtanttagtttgagcgcgcgaigtgcgcaatgcacagacataiygcctatgcgatt	759					
Dd	721	GGCACAGTCAACTTGTGATGTCGAAAAAATTTAAACATACCCCATATTGTTTTTACCAACA	780					
OY	760	gaaggctggcagcagacagcagcgtttcggcaatluttittlatgcaaaaaaacaataaa	819					
Dd	781	GAGGCTATTTTGGCGACAGTGGCTACACCAATGTTTATCTATGTATGCCCAAACCAAAAG	840					
OY	820	tgagctctcaacgagttltaacaaacgcgcctacacctlatlcgcgcagggaaaagcgttc	879					
Dd	841	TGTTTAATTAATTTGGGTATTGCAAAACGGGCAACCCCTATATAGSAAAAAGCAATAGCGTTC	900					
OY	880	cagcgtacaacgcaaagatigtttctaagatgaacatttacagaggcgatacacatacgtc	939					
Dd	901	CAGCGGTTCTTAAGATTTGGTTTATGATCAAAATCTTTGCTGGAGATTCATTCAGTA	960					
OY	940	tttttgaacgcgcagtiaacgacatttttccttcatccaacaaacacgcgtlacggt	999					

Db 961 TTCTACGAACCGTCAAAATGGGAATACTCTTTTAAACGACGTAAATATGGCACAGCA 1020
Oy 1000 acggtacagaacccaacgaaagintlccaatccaaa---gcttaagtaacagacagtc 1056
Db 1021 AAATGCAATGCCAAGACANTGAAACAATTTCTGCTTAATGATTAAAAACGAAACGGT 1080
Oy 1057 cgaacttgaagaaactcttgaatgaactgataaagaaccagtt---acggcgacagg 1113
Db 1081 CAATTCTTAATGTTTCTTTATCCGAGACAGCAAGAACCTGTTTATCATGCTGCAGGT 1140
Oy 1114 ggtgttaacagtaacggtcccaaggttaacaacggtgaaaaccccttccatcagttac 1173
Db 1141 GGTGTCAACAGTTATGACACCCAGACTCAATATGAGAGAAAAATTTCTTTATTGACGAA 1200
Oy 1174 ggcacgycgaactcaltctatacaacaacatcaaccaagcgcgcggttctgtattt 1233
Db 1201 GGAAGGCGAATTGATTAATTACAGCAACATCAATCAAGGTGCTGAGATTATATTTTC 1260
Oy 1234 gaagtgatttcaagtcgtctcgctgaaacaacgaaacgtgacgaagcgcggttcat 1293
Db 1261 CAAGGAGATTATTCGCTGCTGCTGAAATTAACGAACCTTGCAAGGCGCGGCGTTTCA 1320
Oy 1294 atcagtgaaagacagtaacggttacttgaaaglaaacggtgacgaacgacgctgtcc 1353
Db 1321 ATCAGTCAACAGTACCTGCTTACTTGGAAAGTAAAGCGGTGGCAAGACCGCTGTCC 1380
Oy 1354 aaaaatcgcaaaagcagcgtgacggttcaagccaaggggaaaccaaagctgcatacgc 1413
Db 1381 AAATCGCGCAAGGCAAGCGCTGCACGTTCAAGCCAAAGGGGAAACCAAGGCTCGATCAGC 1440
Oy 1414 gtggcgacaggtcagtcatttggatcagcagcgacagcgaataaaggcaaaaacaagcc 1473
Db 1441 GTGGGCAACGGTACAGTCATTTTGGATCAGCAGCGACGATAAAGGCAAAAAACAAGCC 1500
Oy 1474 tttaagtaaatcgcgtcgtntcagcgcgaggggtacggtgcacgtgaatgcgataatcag 1533
Db 1501 TTTAGTGAATCGGCTTGCTGTCAGCGCGAGGGGTACGGTGAACCTGAATGCCATATCAG 1560
Oy 1534 ttcaaccccgcaaaactctattcggtcttcgcgcgagcgtttgatttaacgggcat 1593
Db 1561 TTCAACCCCGCAAACTCTATTTCGGCTTTCGGCGGACGTTGATTAAACGGGCAAT 1620
Oy 1594 tgcgttcgtccacacgtatccaataacgaatgaagggcgatgatcgtncatccataat 1653
Db 1621 TCGCTTTGCTTCCACCGATTCAAAATACCGATGAAGGGCGATGATGTCAACCCAAAT 1680
Oy 1654 gccacaacaacatccacggttacatlaacagggaaatgaagtatacacacacgagtgat 1713
Db 1681 CAAGACAAGAAATCCACCGTTACCATACAGGCAATAAAGATATTGCTACAAACCGG---- 1736
Oy 1714 aagaataatcaatgaactaattacagcaagaagaattgcttacaagaagttggtttcgag 1773
Db 1737 --CAATTAACAACACTTGGATTAGCAAAAAAATAATGGCTTACAGCGTTGGCTGGCAG 1794
Oy 1774 aaaaatacagcaaaacgaacggtcggtcacaactgtttacacgcccgcgcaagaagac 1833
Db 1795 AAAAGTACGACCAAAACGAACGGGGGCTCAACCTTGTTTACACGCCCCCGAGAAAGAC 1854
Oy 1834 cgcacccmgctgtcttcgcgcggaacaaatttaacgycgaacatcagcaacaacagcg 1893
Db 1855 CGCACCCGCTGCTTTCGGGGGAACAATTTAAACGGCAACATCCAGCAAAACAAACGGCG 1914
Oy 1894 aaactgttttcaagcggaagcgacacgacgacgtctacatcatctagaagcggttgg 1953
Db 1915 AAACGTTTTCAGGGGAGACCAACCGCACGCTTACAAATCATTTTAAACGACCAATTGG 1974
Oy 1954 tcaaaaatggaagtatacccaagaagagaatcgtgtgtggaacaacgactgtaacacgc 2013
Db 1975 TCGCAAAAAAGGGGCAATTCGTGGGGGAATCGTGTGGACAAACGACGTGGATCAACCGC 2034
Oy 2014 acgtttaaagcgaaatttccatctcaaggcgggcagcggtgtatttcccgcaatgtt 2073
|||||

Db 2035 ACATTTAAAGCGGAACCTTCCAAATTAAGCGGACAGCGGTGTGTTCCCGCAATGTT 2094
Oy 2074 gccaaagtggaaagcgatgtncaattggaaatcaacgcccgaacagtttgggtgcga 2133
Db 2095 GCCAAAGTAAAGCGATTTGGCAATTTGAGCAATTCACGCCCAAGCAGTTTGGTGTGCA 2154
Oy 2134 ccgcatcaagaacatacaatctgtacaacttgcgaactggaacngtctgcgaacattgtgtc 2193
Db 2155 CCGCATCAAAAGCCACACAAATCTGTACAGCTTGGAGCTGGAGGGGTGTGACAAATTTGTCTC 2214
Oy 2194 gaanaaancattacagcagaataagtgatctcattactgaagaacngatnaagcgcc 2253
Db 2215 GAAAAAACCATTCACGACGATTAAGTGATGCTTCATTGACTAAGACCGCATACGCGGC 2274
Oy 2254 antgtgaagcttmcnathcgnthntlnaanoctmcnnggcnthgmcnacthaungc 2313
Db 2275 AATGTGATCTTCCGCTACGCTCATTTTAATTTCAAGGCGCTTGGCACACTCAACGGC 2334
Oy 2314 aatccttagtgcaatgycgaatacaggtlatacagtlcagccacaacgcccacaaacggc 2373
Db 2335 AATCTTAGTCAATGCGCATACAGCTTATACAGTCAGCCACACGCGCACCAAAACGGC 2394
Oy 2374 aacctagcctcgttgggcaatgcccagaacacatlttaacgaagccaatlaaagcgacac 2433
Db 2395 AACCTTAGCCTCGTGGCAATGCCCAAGCAACATTTAATCAAGCCACATTTAAACGGCAAC 2454
Oy 2434 ncatcgntcggcgcaatgcttcaatctaaagcaacaacgcccgaacaaacggcagt 2493
Db 2455 ACATCGGCTTGGGCAATGCTTCAATTAATCAAGCGACACGCGCTGACAAACGGCGAGT 2514
Oy 2494 ctgacgcttccgcaacgctlaaagcaaacglaagccattccgcactaacgycgaatgc 2553
Db 2515 CTGACGCTTCCGCAACGCTAAGCAACCTAGCCATTCGCGACTCAACGGTAAATGTC 2574
Oy 2554 tccctagcgatgaaggaatgttccatttgaaaaacgacgcttaccggaacactcagc 2613
Db 2575 TCCCTAGCCGATTAAGGAGATTCATTTTGAAGCAGCCCGCTTACGGCAAAATTCAGC 2634
Oy 2614 ggcagcagaaganaacagatlaactlaaagaacagcgaaatgtagcgtgcgtcaagcag 2673
Db 2635 GCGGCGAAGGATACGGCATTTACCTTAAAGACAGCGAATGAGCGCTCGCGACAGCAAG 2694
Oy 2674 gaattaggaattlaaaccttgacaagcgcaacatlaactcaattccgcctatgcacac 2733
Db 2695 GAATTAGCAATTTAAACCTTGACACGCCACCAATTAACCTAATTCGCGCTATGCGCAC 2754
Oy 2734 gatgctcgaagcgcgcaaacccggcgagngtgtcagacacgcgcgcgctgcg----- 2787
Db 2755 GATCGGCGAAGGGCGCAACCGGCAAGTGCAGACAGATGGCGCGCGCGCTTCGGCGCCT 2814
Oy 2788 ---cgcgctccctattatccgttacaacgccaacttcgylagaaatcccggttcaacagc 2844
Db 2815 TCGGCGGCTTCCATTATTCGTTACACGCGCAACTTCGGTGGATGCCGTTCAACACG 2874
Oy 2845 ctgacggtlaaacggaacatgaaacngtcaaggaacatctcgttlatgttcggaactcttc 2904
Db 2875 CTGACGGTAAACGGCAAAATTAACGCTCAGGGAACATTCGCTTATGTGCGAAACTCTTC 2934
Oy 2905 ggtcagcaagcgcaaaattgaagcttgcggaagttccgaaggaacttacaacttgcg 2964
Db 2935 GGTACCGCGACGACAAATTTGAAGCTGGCGGAAAGTTCCGAAGGCACTTACACCTTGGGG 2994
Oy 2965 gtcaacaatlaccggaacacgctgaagcctgcatactgaatgaaagcgttagtgaaggga 3024
Db 2995 GTCAACAATACCGGCAAGCAACCTCAAGCTCGAACAATTTAGCGTATGGAAGGAAAA 3054
Oy 3025 gacacaaacacgctgtccgaaacacttcaattccaccctgcgcaaaacgaacacgttcagtcgc 3084
Db 3055 GACAACAACCGCTGTCCGAAACCTTAATTTACCTTCGCAAAACGACACGTCGATGCC 3114
Oy 3085 ggcggtggttaccactcatccgcaaaagcggcgaattccgcctgcataatccgctc 3144
Db 3115 GCGCGTGGCTTACCACTCATCCGCAAAAGAGCGGAGATTTCGCGCTGCATATATCGGTC 3174
|||||

[illegible]

QY	4225	atgagcgaatggggcgctaaacgcgcgaatcaaaaggttcaocgtctcmtccaagctgcc	4284
Db	4249	atggcgcgatggggcgctaaacgcgcgaatcaaaaggttcaocgtctcmtccaagctgcc	4308
QY	4285	gccgcgcaaaagmccgcgaacttgaagcgcaacacacgcgcygagcatcaaataggctaccgc	4344
Db	4309	gccgcgcaaaagmccgcgaacttgaagcgcaacacacgcgcygagcatcaaataggctaccgc	4368
QY	4345	tggtaa	4350
Db	4369	tggtaa	4374

FEATURES	REFERENCE	AUTHORS	TITLE	JOURNAL
LOCUS	AX044007		4374 bp	DNA
DEFINITION	Sequence 86 from Patent WO0066791.		Linear	PAT 24-NOV-2000
ACCESSION	AX044007			
VERSION	AX044007.1		GI:11342893	
KEYWORDS				
SOURCE				
ORGANISM				
	Neisseria meningitidis.			
	Neisseria meningitidis			
	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;			
	Neisseria.			
	1 (bases 1 to 4374)			
	Piazza,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C.,			
	Masignani,V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M.,			
	Scarlato,V., Rappuoli,R., Frazer,C.M. and Grandi,G.			
	Neisseria genomic sequences and methods of their use			
	Patent: WO 0066791-A 86 09-NOV-2000;			
	CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)			
	Location/Qualifiers			

```

/organism="Neisseria meningitidis"
/db.xref="taxon:487"
BASE COUNT      1266 a      1163 c      1066 g      879 t
ORIGIN

```

Query Match	80.6%;	Score 3508;	DB 6;	Length 4374;
Best Local Similarity	88.4%;	Pred. NO. 0;		

QY	1	atgaaacacacgcgcacaaacgcgaacacacacgcgaagcccgaaacacgcgcgc	60
Db	1	ATGAAACACACGCACAAACGCACAACCGAAACACACCGCAAAAGCCCGAAAAACCGGC	60
QY	61	atccgcctctcgctgcgtacttaagcaatgcctgtcgttcgcgattcttcccaagct	120
Db	61	ATCCGCTTCTCGCTGCTTACTTAAGCCATATAGCCTGTGCTTGGCATATTTCCCAACCC	120
QY	121	tggcgaggacacattatcttcgcgcatcaactaccatactacgcgacttgcggaat	180
Db	121	TGGCGGAGACACTTATTTGGCATCACTACCAATATCTTCCGCATTGGCGAAAT	180
QY	181	aaagcgaagttgcagtcgcggcggaagatattgaggtntacaacaaaaaagggaattg	240
Db	181	AAAGGCAAGTTTGCAGTCGCGGCGCAAAAGATTTAGGTTTACAACAAAAAGGGAGTTG	240
QY	241	gtcgccaatcaatgacaaaaagcccgatgatatttctgtgtgtcgtgaacg	300
Db	241	GTGCGCAATCAATGACAAAAAGCCCGATGATTTCTTGTGTGTGCGGTAAACGGC	300
QY	301	gtcgcgcatctgtgtgcgcgaatcaatatctgaagcgtgcgaacaaacgcgcgtataac	360
Db	301	GTCGCGCATCTGTGTGCGCGATCAATATATTTGTAGCCTGTGCACATAACGGGGCTATAAC	360
QY	361	aacgtgatcttgtgtgcgaagaagaaatcccgatcaagcagcgttcttaccaat	420
Db	361	AACGTGATCTTGTGTGCGAAGAGAATCCGATCAACATCGTTTACTATAAATTT	420
QY	421	gtgaaagaataatataaagcctgcacattca--caccttcaacagcgcgtatncaat	477
Db	421	GTGAAAGAATAATATAAAGCCTGCACATTCA--CACCTTCAACAGCGCGTATNCAAT	477

Db 421 GTGAACGGAATAATTATAAAGCAGGACTAAAGCCATCCTTATGGCGCGATTATCAT 480
OY 478 atgcgcgttttgtataattttgtcacagatgacgaacctgtgcgaatagcagatgacatg 537
Db 481 ATGCCCGCTTTGATAAATTTGTACAGATGACAGAACCTGTGTAATATCACCACTATATATG 540
OY 538 agggggaataccattccgcaataaagaataatcccgacgctgcgcgatccgctcaga 597
Db 541 GATGGCGGGAATATATACGATCAAAAATATTACCTTGACCGCTGTTCGATATGGGACAGC 600
OY 598 caccactatgscgtatgatagtatga-----caacaacggcgatlttaactcactcgcgc 651
Db 601 AGGCAATATTTGGGATGATGAGATGAGCCCAATAAACCGCAAAAGTTTCATATCATATT 660
OY 652 gca-----tggttaattggcgacatcacatagcaggtgtggggaataat 699
Db 661 GCAAGTGGCTATTCTTGGCTGTTGGGCAATACCTTTTGACAAAATGATCAGTGTGT 720
OY 700 ggcgtanttagtttgaacgcgcatgtgcgcatagcacaagctatgscctatgcgcat 759
Db 721 GGCACACGTCAACTTAGTATGAAAAAATTAAACATAGCCCATATGCTTTTACCACACA 780
OY 760 gcaagtgccgacgagcagcagcgttcgcgaattgttatatttgcacaaaacaataaa 819
Db 781 GAGAGCTCATTTGGCGACAGTGGCTCACCAATGTTTATCTATGATGCCCAAAAGCAAG 840
OY 820 tggcgtgcacagagttttaaacaacgcgctacacttaatccgcaaggaaacggcttc 879
Db 841 TGGTTAATTAATGGGGATTGGCAAAAGGGCAACCCCTATATATGAAAAAGCAATGGCTTC 900
OY 880 cactctgaacgcaaaagtattgtctacagatgacattacagagcgacacacatacgcgc 939
Db 901 CACGTGTTGCTAAGATATGTTGTTATGATGAATCTTTGGCGAGATACCATTCACATGA 960
OY 940 tntttgaaccgcgacgtacagacattttcccttaccatccacaacaagcgatcggt 999
Db 961 TTCTACCAACCACTGTCAAAATGAGAAATACCTTTTAAAGAGATATATATGGCACAGGA 1020
OY 1000 jacggtacagaacacgaacgaagaagtttccaatccaac---gcttaagttacagacgtc 1056
Db 1021 AAAATCATATGCGCAACATATGACACACATTTCTGCTTATACATTTAAAAACAGAACCTTT 1080
OY 1057 cgaactgttgaacgaactcttgaatgaacatgataagaacgaactt---acgcgcgcaggg 1113
Db 1081 CAATGTTTATATGTTCTTTATTCGAGACAGACAAAGAACCTGTTATATCATCTGAGGT 1140
OY 1114 ggtgttaatcagtaacggtccaaggttaacaacggttgaacaccttcttaccgtatc 1173
Db 1141 GGTGTCAACAGTTATWCGCCAGACTGAATATATGAGAAATATTTCCCTTTATTTGACGAA 1200
OY 1174 ggcgaacggcaaacatcattcttaacaacacatcaacgaagcgcggtgttgaattt 1233
Db 1201 GGAAGAAGCGCAATTAATTAACCAACATCAATCAAGGCTGTGGAGATATATATTTTC 1260
OY 1234 gaaggtatlttlaacggttcgccttgaacaaacaacgaacggttcgaagcgcggttcac 1293
Db 1261 CAAGAGATTTTACGGTTCGCTGSAATAAATACGAACCTTTGCAAGGCGCGGCGTTCAAT 1320
OY 1294 atcagttgaagacagtlaccgcttaacttgaagaagtaaacgcgtgtgcgaacgacgcgtctcc 1353
Db 1321 ATCAGTGAAGACAGTACCGTTACTTGGAAAGTAAACGGCGTGGCAAAACACCGCCGTGTC 1380
OY 1354 aaatctggcaaaagcagctcagcttcaagtcgaaggggaacaaacgaagctcagatcagc 1413
Db 1381 AAAATCGCAAAAGGACGCTGACGTTCAAGCCAAAGGGGAACCAAGGCTTGATTCAGC 1440
OY 1414 gtggcgacggttacagttatlttgatcagcagcgacagacataaagcgcaaaaacaacagc 1473
Db 1441 GTGGCGGACGATAGCTATTTTGGATCAGCAGCGACGAGATTAAGGCAAAAACAAAGCC 1500
OY 1474 tttagtgaatcgcgttntcaagcgcgaggttacggtgcacatgaaatgcgcatatcag 1533
Db 1501 TTATGTGAATTCGGCTTGTCAGCGCAGGGGTAGCTGAACCTGAATGCCGATATATTCAG 1560

OY 1534 ttcaaccccgacaacactctatlttcgcgttctcgcgcgcagcgttttgaatttaaacgggcat 1593
Db 1561 TTCAACCCCGACAAACTCTATTTGCGCTTTCGGCGGCGAGCTTTGGATTTAAACGGGCAT 1620
OY 1594 tgcgttcttccacgcgtalttcaaaataccgaatgaagggcgatgtatgcnatcaat 1653
Db 1621 TCGCTTTCGTTCCACCGTATTAATAATCCGATTAAGGGGCGATGATTTGTCMAACCATAT 1680
OY 1654 gccacaacaacatccacgtttaccatattacagggaaatgaagatattacaacacggagtgt 1713
Db 1681 CAAGACAAGAAATCCACCGTTTACCATTACAGGCATTAAGATATTGCTACAAACGG --- 1736
OY 1714 aagaatcatcagacttaatttcaagaacgaatltgcctcaacaggtttggtttgcgag 1773
Db 1737 --CAATTAACAACAGCTTTGATAGCAAAAAGAAATTTGCTTACACAGGTGTTGGCGAG 1794
OY 1774 aaagatacgaaccaaaacgaacggcggtcgaacctgtlttaccagccgcgcgaagaac 1833
Db 1795 AAAGATATGACACCAAAACGAACGGCGGCTCAACCTTGTATACGACCGCGCGAGAAAGC 1854
OY 1834 ggcacccngtgcgttctccgcggaacaaatttgaacgcaacatcagcacaacaacggc 1893
Db 1855 GCGACCCCTGCTGCTTTCGGCGGACAAATTTAAACGGCAACATCAGCAAAACAAACGGC 1914
OY 1894 aaactgttlttcaagcgacagacacgcgcgccttaacatlttgaagcgggttgg 1953
Db 1915 AAACGTGTTTACGGCGGACAGCAACGCCGCTCAATCATATTAACGACCATTTGG 1974
OY 1954 tcaaaaatggaaggtatcccaacgaagaatcgtgtgagcaacagatgatcmaacgc 2013
Db 1975 TCCCAAAAAGAGGCAATTCCTTCGCGGGGAATTCGTGGGCAACAGATGATCAACCGC 2034
OY 2014 acgtttaaagcgggaatltccatcatcgaagcgcgacggcggtatltccgcgaatgt 2073
Db 2035 ACATTTAAAGCGGAAAACTTCAATTAATTAAGCGGACAGCGGCGTGTTCGCGCATGTT 2094
OY 2074 gccaaagtgaaggtatcgnacatttgaacaaatcaacgcgcgaagaagtttgggtgcga 2133
Db 2095 GCCAAAGTGAAGGCGATTTGGCATTTGAGCAATACGCCCAAGCGATTTTGGTGCACA 2154
OY 2134 ccgcataaagccatacaatctgtacacggttcgcgacttgaacngtcttgaacaaatgtgtc 2193
Db 2155 CCGCATCAAAAGCCACACATCTGTACAGTTCGCTGAGCGGCTGTACAAATTTGTGTC 2214
OY 2194 gaanaaancatltaccgcgataaagtgtgtcttcatgtactgaacgaacgaacgtc 2253
Db 2215 GAAAAAACCATTAACGACGATTAAGTGTATTCCTTCATTGACTAAGACCGACATCAAGCGC 2274
OY 2254 atgtlnagctlnacnaatnaacnlnatnaaancnlnacnngcgtlncmactnaangc 2313
Db 2275 AATGTGATCTTGGCGATACGCTCATTTAATGTACAGAGGCTTGGCACACTCAAGGCG 2334
OY 2314 aatcttagtgaatgtgcgatacagcgttatatcagtcagccacaacgcaccccaaaacggc 2373
Db 2335 AATCTTGTATGCAAAATGGGATACAGTATACAGTACACCAAAAGCCACCAAAAAGGCG 2394
OY 2374 aaccttagccctgtgggaatgcccacaacatlttaacgaagccacatlttaacggcgaac 2433
Db 2395 AACCTTAGCCCTGTGGGCAATGCGCCACCAACATTTAATCAAGCCACATTTAAACGGCAAC 2454
OY 2434 ncatcgnnttccgcaatgtcttcaatcgaacaaacacgcgcgcacaacaaacggcagc 2493
Db 2455 ACATCGGCTTGGGCAATGCTTCAATTTAATGTAAACGACACAGCGGTACAAAAGCGCAGT 2514
OY 2494 ctgacgcttccgacaacgctaaaggcaacgtaagcgaatltccgactcaacggcaatgtc 2553
Db 2515 CTGACGCTTTCGGCAACGCTTAAGCAACGTAACGTAATTCGCACTCAACCGTATATGTC 2574
OY 2554 tcccttagcggatgaagcgtatttccatttgaacacgcgcgtttaccgcgaacactcagc 2613
Db 2575 TCCCTTAGCCGATTAAGGCGATTTTCATTTTGAAGACGCCGCTTTACCGGACAAATTCAGC 2634

OY	2614	ggaagcaagaaanacagattactacttaaaagacagcgaaaygacgtgycgttcaggaacg	2617
Db	2635	ggggcgcaaggatnccggcatttaccatttAAAAGACGCAAAATGAGCGTCGCCGTACGGGCACG	2694
OY	2674	gaattaggcaatttaaacctcttgacaacgccaacttaacatcaattccgcctatocgcaac	2733
Db	2695	GAATTAGGCAATTAAACCTTGACAAACGCCACCTTACACTCAATTCGGCTATCGCCAC	2755
OY	2734	gagctcgcagcgccgcaaacccgcaagtgctcagacaacgccgcgcgcgttcg-----	2787
Db	2755	GATGCGCGCAGGGCGCAACCGGCGACAGGCAACACATCGCGCGCCCGCGCTTATCGCGCGCT	2814
OY	2788	---cgccgtccccattatctcgttcaacccgcaacttcgttagaatccggttttaacgcg	2844
Db	2815	TCGCGCGCTTCCCTATTATCCGTTACGCCGCCAACCTTCGGTAGAATCCGTTTAAACACG	2874
OY	2845	ctgacggtataacggcgcaattgaaacngtcaaggaacattccgcttattgtcggaaactctc	2904
Db	2875	CTGACCGGTAAACCGGCAAAATTGAACGGGCGACGAGAAACATTCCGCTTATGTGCGGAACTCTTC	2934
OY	2905	ggctaccggaagcgacaacaaattgaaagctcggcggaagagttccggaaginatctaaccttgycg	2964
Db	2935	GGCTACCGGACGACCAAAATTGAAAGCTGCGGGAAGTTCGAAGGCACTTACACTTGCGCG	2994
OY	2965	gtcaacaataccgycgaacgaaccgcgtaagcgtcatcatttgacgcyagtgaygaaggaaa	3024
Db	2995	GTCAACAATAACCGGCNAACGNAACCTCGAAGCCTCGAACAATTTGAACGGTAGTGGAGGAAAA	3054
OY	3025	gacaaacaaaccgctgtgcggaanaaccttaattcaacctgcgaanaacgaaacgctcatgccc	3084
Db	3055	GACACAACAAACCGCTGTGCGGAACCTTTAAATTACCTCGTAACCAACGATCGTAGTGCC	3114
OY	3085	ggcgcgtgcgcttaccacactcatccgcgaagacgcgagctccgcctgcataactcgcgttc	3144
Db	3115	GGCGCGTGCCTTACCACACTCATCCGCAAAAGACGGGAGTTCGGCTGCATATCCGCGTC	3174
OY	3145	aaagaaacaagagcttccgcgaacaactcgcgcaagcgcaagccaaaacagcgcgaaanaa	3204
Db	3175	AAAGAACAAAGACTCTTCCGCAACAACTCGGCAAGGCAAGAACCAAAAACAGGGCGGAAAA	3234
OY	3205	gacaaacgcgcaaaagcctltgacgcgctgatatgcgagccgagcgatgctgcgcggaanaaga	3264
Db	3235	GACAAACGCCAAAGCCTTGACCGCGGTGATGTGGCGCGCGCGCATGCCGTGAAAACACA	3294
OY	3265	gaaagcgtttgcggaacccgycgcgcmgacgycgaggaanaatgttcggcatlatacgagcg	3324
Db	3295	GAACCGTGTGCGGAACCGGCGCGGACGAGCGGGGGAATTCGCGCATTTATGCAAGCG	3354
OY	3325	gaggaagagaanaaacggtgtcgaagcggaataaagacgcmntltgacgaacagcgagaa	3384
Db	3355	GAGGAAGAAGAAAAACGGGTGCAGGGGGAATAAGACACCGCCTTGGCGGAACACAGCGGAA	3414
OY	3385	gcggaaccccgccggnltaccacacgcgccttcgccgcgcgcgcmgagcgccgcgcggaatttg	3444
Db	3415	GCGGAAACCCCGCGCGCTACCAACCGGCTTCCCGCGCGCCGCGCGCGCGCGGAAATTGG	3474
OY	3445	ccggaaccccgcaagccccaacacgcaactcaaaccccaacccgagcgagcactgtatnagccgt	3504
Db	3475	CCGCAACTGCAACCCCAACCG-----CAGCCCCAAACCGAGGCGCACTGTATCAGCCGT	3528
OY	3505	tatgcgaataagaggtttagtgaatttccgcgaacgctcaacagcgttctgcgctgaag	3564
Db	3529	TATGCCAATTAGCGGTTTGATGTAATTTCGCGACGCTCAACACGCTTTTCGCCGCTACAG	3588
OY	3565	gaagcaattgacccggtgttttcgcgaagacgcgcgcgaacgcmgttggacaagcmgcatc	3624
Db	3589	GACGAATTAAGACCGGTATTTCGCCAAGACCCCGCAACGCCGTTTGGACAACCGGCATC	3648
OY	3625	cgnaacacaaacacatacgttccgcaagaattccgcgctaccgcaacaaacacccgcgtg	3684
Db	3649	CGGACACCAAAACATACTACCGTTCCGAAGATTTCGCGGCTTACCGCCAAACAAACCGACTTG	3708
OY	3665	cgcgaacatcgatgtgcagaaaaaacctcgcgcagcgcgcgcgtcgcgaatccgttlttcgcac	3744

Db	3709	CGCGAATCGGTATGCGAAGAAAAACCTCGGCGAGCGGCGGTGCGATCCGTGTTTGCAC	3768
QY	3745	aaccgacccgaaaaacncttcgacgacgcatcgtgaactcggcaagcttgcacacg	3804
Db	3769	MACCGAACCGAANAACCTTCGACGACGGCATCGCACTCGGCAAGCTTGCCACGGC	3828
QY	3805	gccgtttcgggcaatacggcatcggcaggttcgaatcgtgacatcgaacacgagcgcggt	3864
Db	3829	GCGGTTTTCGGGCAATAACGGCATCGACAGTTCATCATCGCATCGACGCGGGCGGGT	3888
QY	3865	tttagaagcgagantctlnlcagacggcatcggaggaataatccgcgcggtgtctgcat	3924
Db	3889	TTTAGAGGAGGCGACCTTTCGACGCGCATCGAGGCAAAATCCGCCGCGCTGCTGCAAT	3948
QY	3925	tacggatctcaggaacgatacgcgcggtttcggcggaattcggacatcgaacgtatc	3984
Db	3949	TACGGCATTCAGGACAGATACCGCGCGGCTTTCGGCGGATTCGGCATTCGACCGCATAC	4008
QY	3985	gagcgcaacgagctattcgtccaaaaagcgattacgcgacgacgaacgltcaatatcgcc	4044
Db	4009	GGGCGAACCGGCTATTCCTCCAAAAGCGGATTCACCGCTACGAAACGTCATATTCGCG	4068
QY	4045	accccgagcttcggtcgtcaacgntacgngagggcgatgaagaagcatattcatc	4104
Db	4069	ACCCCGCGGCTTCGATTCACACCGCTACCGCGCGGCGCATTAAGCGAGATTAATTCATA	4128
QY	4105	ccggcgcaacacatnltcaatcanccttattnaagcgtgtcctatacgaatgcggttcg	4164
Db	4129	CCGGCGCAACACATTCATTCATCAGCGCTTATTTGAGCCTGTCTCTATACGATCGCCTTCG	4188
QY	4165	ggcgaagtcggagacgcggtcaatccgcnfatttggtcctagaatttcggaaaaacccgc	4224
Db	4189	GGCAAAAGTCGAGACACGCGTCAATACCGCGATTATGGCTCAGAGATTTCGGCAAAACCGCG	4248
QY	4225	agtgcggaatggggcgataacgcgcgaaatcaaaaggtttcaagctgttcnctcagcgtgc	4284
Db	4249	ACTGCGGATATGGGCGTAAACCCGCAAAATCAAAAGCTTTCACGCTGTCCCTCCACGCTGC	4308
QY	4285	ggcgcaaaagggcgcaactggaagcggaacgacgaggggataaatttggcgtacgc	4344
Db	4309	GCGCGCAAAAGCGCCCGCACTGGAAGCGCACACAGCGGGGCATCAAAATTAAGCTACGCG	4368
QY	4345	tggttaa 4350	
Db	4369	TGGTAA 4374	
RESULT 6	AE002547/c		
LOCUS	AE002547	10074 bp	DNA linear BCT 25-MAY-2000
DEFINITION	Neisseria meningitidis serogroup B strain MC58 section 189 of 206		
ACCESSION	AE002547	AE002098	
VERSION	AE002547.1	GI:7227244	
KEYWORDS			
SOURCE	Neisseria meningitidis MC58.		
ORGANISM	Neisseria meningitidis MC58		
REFERENCE	1 (bases 1 to 10074)		
AUTHORS	Tetelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwinn, M.L., Deboy, R., Peterson, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E., Clifton, H., Clark, E.B., Cotton, M.D., Uitterlidge, T.R., Khouri, H., Olin, H., Vamthepan, J., Gill, J., Scalapato, V., Masignani, V., Pizsa, M., Grandi, G., Sun, J., Smith, H.O., Fraser, C.M., Moxon, E.R., Rappuoli, R. and Venter, J.C.		
TITLE	Complete genome sequence of Neisseria meningitidis serogroup B strain MC58		
JOURNAL	Science 287 (5459), 1809-1815 (2000)		

MEDLINE
2015755
10710307
2 (bases 1 to 10074)
Tetelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eissen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwin, M.L., Deboy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischman, R.D.,
Dougherty, B.A., Mason, T., Cleecko, A., Parksey, D.S., Blair, E.,
Cltone, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H.,
Qin, H., Yamathavan, J., Gill, J., Scariato, V., Maignan, V.,
Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappuoli, R. and Venter, J.C.
Direct Submission
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
Location/Qualifiers
1. 10074
/organism="Neisseria meningitidis MC58"
/strain="MC58"
/db_xref="taxon:122586"
/note="serogroup: B"
704. 1375
/gene="NMB1983"
704. 1375
/gene="NMB1983"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42311.1"
/db_xref="GI:7227245"
/translation="MPPDKYKLSKKNIPFAKYVAQIHNLSPENCOTLLAQTEOI
HKKNAVASLEDIOTLLNKRAYQVYVSHISNCEPDVDSITKINNIYAKDLSLARS
DRTGSGVATLLDGSRAHPNPKVEIYAVRNLGLOSSTTEAVALPMYCMQOQV
MGCKRTATLFAAGLMMAGCGCILEISEMQMPQFNEKLSAFYRSGDDTDSKIFYQNC
ISGID"
1437. 2548
/gene="NMB1984"
/note="This region contains an authentic frame shift and
is not the result of a sequencing artifact; similar to
SP:000840 GB:211857 PID:45066 percent identity: 97.40;
identified by sequence similarity: putative:IS1106
transposase, authentic frameshift"
complement(2864. 7237)
/gene="NMB1985"
complement(2864. 7237)
/note="similar to GP:1408569 percent identity: 76.08;
identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="adhesion and penetration protein"
/protein_id="AAF42312.1"
/db_xref="GI:7227246"
/translation="MKTDRKRTFTHHKAPKTRGRSPAYLAICLSFGLPQAMAGH
TYGIVQYVRFAENKGRFAGVAKDIENVKQGLVSKMTAKPMIDSVSRGVA
ALVGDQYIVAHNGGKNNYDFAEGNPNPOHREPTKYKNNYKAGTGHYGGDYH
MPRLHFTVDAEVEMTYMDGRKTIIDNNTPDVRVACAGQYMRSDDEPRNNRSY
HISAYSMVGVGNTFAONGSGGTVNLGSEIKISPGFELTGSFGSGSWMFYDA
QOKWMLNGVLOTGNPYIGKNGFQVLAKDEYDEIFAGDLSVFEYERONKYSFND
DNNGTGKLNKHEHNSLPNRLKRTVOLFNLSLETAEPYHAAGVNSYRPLNG
ENISFTDEGKELLTNSINOGAGGLYFOGDFTVSPENNETWOGGVAHISESTYWK
VGVANDRLSKIGKGLHNOAGKNGOSISYDGTLLIDQADDKCKROAFSEITLVS
GRGTVOLNADNGRPDKLYFGRCGRDLNKHSHLSFPHQNTDEGAMVYNNHODKEST
VTTGNKDIAFTGNNNSLDSKKEITAYNGWEGEKDTKTNGRLNLTVOQPAEDRTLLS
GNTNLNNGITQTNKLEFSGRPPTHAYNHLMDHNSQKGIPIRGELIWNWDYLNRFKA
ENFOIKGCAVAVSHNAVKVGDMLSHNAQAVFGVAPQSHITICRSDMTGLTNCVK
TITDDKVIASLTJTDISGNDLADHAHLNGLATLNGLNSANGDRTYVSHNATNG
NLSLVGAQAFENQATLNGNTSASGNASFNLSDAHVOGSLTSGNAKAVNSALNG
NVSILADKAVHFESSRFTGQISGCKDRALHDKDSWPTLPSGTGSLNLDNAITTLNS
AVRHDAAGTGSATDAPRRRSRSRSLSVTPPTSPESRNTLTNGKLNGOQTR
FMSLFGYRSDKTLAESSEGTYTLAVANNTPESLTQTLVVECKDNKPSENLFT

LONEHVDAQAMRYOLIRKDFEHLNPKVEQELSDKLCKAEAKKQAEKDQASLALI
AAGRDVAKTESVAEPARQAGENGNGIMQAEDEKRYQADKDTALAKQAEETRPWT
AFPRARRARLDLPOLPOPOPOPODRLISRYANSGLSEFSAITLNSVFAODELDTVFA
EDRRNAVWTSCTIRKHYRQODFRAYROOTDRLQIGMOKNLGSRGILFHSNTEYNT
FDDGIGNSARLAKHCAVFCQYCDIRYVITISACAGSSGSLSGIGCKIRRYVHYGIO
ARYRAGFEGGFEIPIHIGATRTFVQADRTREVNATPELARNRYRAGIKADYSPFA
OHISITPYLSYSTYDPAASGKVTFRVATLADDFGKTRSAEWGVNAIEGFTLSLHAA
AAKGPOLEAQSAGIKLYRW"
complement(7371. 7802)
/gene="NMB1986"
complement(7371. 7802)
/gene="NMB1986"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42313.1"
/db_xref="GI:7227247"
/translation="MTENADKARQAVETVKSPELVEQILSDEYVQIMARCFHSGP
LPPSDLAQYNDIISGADRIMAMAEKQDAVHETIRODOTFNRQQLYGFISAILIL
LFAVFLVSGYPATASLAGTVPFALAGAFVIGSRHDOGN"
8087. 9433
/gene="NMB1987"
8087. 9433
/gene="NMB1987"
/note="similar to GB:L10328 SP:P25522 PID:290554 GB:U00096
PID:2367268 percent identity: 69.20; identified by
sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="thiophene and furan oxidation protein ThdP"
/protein_id="AAF42314.1"
/db_xref="GI:7227248"
/translation="MSDNYPTLAAVAVTAPGRGGVGYIRISGNILPMAALCKTPEP
RTATYADFTDTGQALDSGLLFFRAPASTGSDVIELOSGGPPYWDMLNRLTEIG
ARLAEPEFTKRAFLNDKLDLQAEVADITDSSNSAHLALRSKGDSSRTHGIV
EDLITRLMVEATLIDPREDIDFLEADADKLDGLRRADVLANAQQAALIREGLN
VVLVGAIPNGKSLNALAGDEVAIYTDIAGTTRDAVREBILIDGVVNHVYDAGLE
TDDVYRIGIERSRAVESADVALVDPREGLNERTRAILIDALPELKRLEIHSKSD
LHAHAAGRGCTGAFTYIALSAKGGDGLDKRTLRLRAGWQSGSEGLFARTRHVNL
KAAQDELISALACGNHQTLEFPAHLRLAQAQCEITGEFTADLLTVIFRFLCGK"
BASE COUNT 2307 a 2467 c 2721 g 2579 t
ORIGIN
Query Match 80.6%; Score 3508; DB 1; Length 10074;
Best Local Similarity 88.4%; Pred. No. 0;
Matches 3876; Conservative 0; Mismatches 462; Indels 48; Gaps 8;
QY 1 atgaacaacgcaacgaacgaacgaacgaacgaacgaacgaacgaacgaacgagcgc 60
DB 7237 ATGAAACCAACGCAACGCAACGCAACGCAACGCAACGCAACGCAACGCGCCG 7178
QY 61 atccgcttcgcgcgttacttagccatagcctgctgcgttcgagatcttcccaagct 120
DB 7177 ATCCGCTTCGCGCTGCTTACTTACATATGCGTTCGCGGATTCCTCCCAACGC 7118
QY 121 tggcgagggaacacttcttgcgcatcaactcaactactcggaacttgcggaat 180
DB 7117 TGGCGGGGACACATTATTTCGGCATCACTACCAATCTATCGGACTTTCGCAAAAT 7058
QY 181 aaaggcaagtttgagtcgagtcgaggaagataatgaagntacaacaaagggagttg 240
DB 7057 AAAGGCAAGTTTGACGTCGGCGGCAAGATATTGAGCTTACAAACAAAAGGGAAGTTG 6998
QY 241 gtggcgaatcaatgcaaaaaagcccgatgatttctggtgtcgcgtgaagc 300
DB 6997 GTGGCAAAATCAAAACAAAAGCCCGATGATTTTCTGTGTGTGCGGTAAAGGC 6938
QY 301 gtggcgagcatgtgtggcgatcaatatacttgagcggtgcacataacggcggtataac 360
DB 6937 GTGGCGGCAATTTGATGGCGGATCAATATATTGTGACGCGGCAATACGGCGGTATAAC 6878

DB 4723 CTGACGCTTCCGGCAACGCTAAGCAAAAGCTATTCGCGACATCAACGATATSTC 4664
QY 2554 tccctagccgataagcagatcatccatttgaagaacgcccgtttaccggaacatcagc 2613
DB 4663 TCCCTAGCCGATTAAGCAGATTTCCATTGTGAAGACCGCCTTTACCGGCAATAGC 4604
QY 2614 ggcagcaaganaacagcatcacacttaaaagacagagatgagacgttcgctcagcagc 2673
DB 4603 GCGGCAAGGATTCGGCATTTACCTTTAAAGACACGCCAATGAGCGCTGCCGTACGGCAG 4544
QY 2674 gaattagcaatttaaaccttgaacaacgcaacatlaactcaattccgctatcgccac 2733
DB 4543 GAATTAGGCAATTTAAACCTTTGAACAGCACCATTTACATCAATTCGCTATCGCCAC 4484
QY 2734 gatgtcagagcgcaaacgcaagngtgcagacgcccgcgcgttcg----- 2787
DB 4483 GATGCGGCAAGGGGCGCAAAACCGGCAAGTCGACAGATGCGCGCGCGCTTCCGCGCGT 4424
QY 2788 ----cgccggttccctattatccgtttacacgccaacttcggtagaatccggttcaacagc 2844
DB 4423 TCCGCGCGTTCCTATTTATCCGTTACCGCGCAACCTTGCGTAGAATCCGTTTCAACAGC 4364
QY 2845 ctgacgtaaacgcaaatgaacngtcaagaaacatccgcttatgtcgaacttc 2904
DB 4363 CTGACGCTAAACGGCAATTTGAACGGTCAGGGAACATTTCCGCTTTATGTCGGAATCTTC 4304
QY 2905 ggtaccggaagcgaacaaattgaagctgcggaagttccggaagacttacccttgcg 2964
DB 4303 GCTTACCGCGGCAAAATTGAAGCTGCGGAAGTTCCGAAGGCACTTTACACTTTGGCG 4244
QY 2965 gttcaacaataccggaacagacccgtaagctgataatgaacgtatgtgaagagaa 3024
DB 4243 GTCAACATATCCGGCAACGAACTGCAAGCTCGAACATTTGAGGTAGTGAAGAAAGAAA 4184
QY 3025 gacaacaacgctgttcgcaaaacttaattccacttcgcaaaacgaaacgacgtatgc 3084
DB 4183 GACAACAAACCGTGTCCGAAACCTTAATTTCAACCTCGAAACGAAACAGACGTGATGCC 4124
QY 3085 ggcgctgtgcgttaacaaactcaccgcaagaagcggaagttccgctcgtataatcgatc 3144
DB 4123 GCGCGCGGGGTATCCAACTCATCCGCAAAAGACGCGGATTCGCTCATTAATCCGGTC 4064
QY 3145 aaagaacaagagcttccgcaaaactcgcgaagcgagaagcgaacaaacagcggaacaa 3204
DB 4063 AAAGAACAAAGCTTTCCGCAAACTCGGCAAGGCAAGGCAACCAAAACAGCGGAAAAA 4004
QY 3205 gacaacgcgcaaacgcttgaacgctgtatgtgcgacgagcgatgcgcgcgcaaaaagaca 3264
DB 4003 GACAACGCGCAAGCCTTTGAACGCGCTGATTCGCGCGGCGGATGCGGTGCAAAAAGCA 3944
QY 3265 gaagaagctgcggaacgcccgcgcmggaagcggaagaaatgtcgcgataatgaagcg 3324
DB 3943 GAAAGCGCTTCCGAAACCGCGCGGCAAGCGGCGGGAATGTGCGCATTAATGACGGG 3884
QY 3325 gaagaagagaaaaaggtgacgagataaagaacagcgcmgttggaacacagcgcaaa 3384
DB 3883 GAGGAAAGAAAAAACCGGTGTCAGCGGATTAAGACACCGCTTTGGCGAAACAGCGCGAA 3824
QY 3385 ggcgaacacgagccggnltaacacgcttcccgcgccgcmgncgacgcccgcggaatttg 3444
DB 3823 GCGGAACCCGGCGCGGTACACCGCCTTCCCGCGCGCGCGCGCGCGGGGATTGG 3764
QY 3445 ccggaacgagccccaacgcaacgcaacccaacgcaacgcaacgcaacgcaacgcaacg 3504
DB 3763 CCGGCACTGCAACCCCAACG-----CAGGCCCAACCGCAACCGGACCTGTATCGCCCT 3710
QY 3505 taagcaatagcggttgagtgaaatttcgcaacgctcaacagaggtttgcgcttacag 3564
DB 3709 TATGCCAATAGCGGTTTGAAGTATTTCCGCGACGCTCAACAGGCTTTTCCCGGACAG 3650
QY 3565 gaacgaattgagcgcgtgttttcgcaagaacgcaacgcmgtttgcaaaagcngacac 3624

DB 3649 GACGAATTAGACCGCGTATTTGCCGAAGACCGCCGCAACGCCGTTTGGACAACGGCATC 3590
QY 3625 cggnaacccaacacactaccgttcgcaagatttcgcgcgtcaccgccaacaaacgacctg 3684
DB 3589 CGGACACCAACACTACCGCTTGGCAAGATTTCCGCGCTACCGCCCAACAAACGACCTG 3550
QY 3685 cgcgaatcgtatgcagaaaaaacctcgacgagcggtcgcgttcgacacccgttttcgac 3744
DB 3529 CGCAAAATGATATGCAAAAAACCTGCGCAGCGGGCGGCTGCGCATCTGTTTCCAC 3470
QY 3745 aacggaacgaaaaacnctcgacgaagcgtatcggaactcgcgaacgcttgcccaagc 3804
DB 3469 AACCGGACCGAAAAACCTTTCGACGAGCGCATCGGCAACTCGGACGCTTGCCACGGC 3410
QY 3805 ggcgtttcggaatacagcatcgcaagttcgacacacgcatcagcagcgcgcggt 3864
DB 3409 GCGGTTTGGGCAATTCGCGCATGAGAGTTTCTACATCGGCATACGCGCGCGCTGCAT 3350
QY 3865 ttaagacgagcancntltaacgagcatcgaggaacaaatccgcgcgcgtgtgcac 3924
DB 3349 TTTAGCAGCGGCAAGCCTTTGACAGCGCATCGAGGCAAAATCCGCGCGCTGCTGAT 3290
QY 3925 taagcattcaagcagatacgcgcgcgtttcgcgagatcggaatcgaacgtaacac 3984
DB 3289 TACGCGATTTCAGCAGCATACCGCGCGGCTTTCGCGGATTCGCGCATCGAACCGCATC 3230
QY 3985 ggcgcaagcgctatttcgtccaaaaagcgatctacgctcaggaacagttcaatcgc 4044
DB 3229 GCGGCAACCGCTATTTGCTCAAAAAGCGGATTACCGCTACGAAACGTAAATATCGCC 3170
QY 4045 accccgcttcgttcaacacnltacacnagcgagcaatlaagagatlatltaacaa 4104
DB 3169 ACCCGCGCTTTCATTTCAACCGCTACCGCGGCGGCAATTAAGCAGATTAATTAATAA 3110
QY 4105 ccgagcaacaacatnltcaatcancctatttaagctgtcctatcagatgacgcttcg 4164
DB 3109 CCGCGCAACACATTTTCATCACGCGCTTATTTGAGCCTGTTCATTCGAGCGCGCTTGC 3050
QY 4165 ggcgaagtcggaacacgagcgtcaatcagcngtatgtgtcagatlttcgcaaaacccg 4224
DB 3049 GGCAAAGTCCGAACACCGCTCATATCCGCGGTATTTGGCTCAGATTTTCGCAAAACCGC 2990
QY 4225 agtgcgaatlggagcgtlaaacgcccgaatcaaaagtttcaacgctgtcncatcagctgc 4284
DB 2989 AGTGCGAATGGGCGGTAACCGCGCAATCAAAAGTTTACAGCTGTCCTCACGCTGCC 2930
QY 4285 gccgcaaaagncgcgaactgggaagcgcaacacgagcggtgacatcaattgctacgcg 4344
DB 2929 GCGGCCAAAGCCCGCAACTGGAAGCGCAACACAGCGCGGCGCATCAAAATTAGCTACCGC 2870
QY 4345 tggtaa 4350
DB 2869 TGGTAA 2864

RESULT 7
AX044034/c 34980 bp DNA Linear PAT 24-NOV-2000
LOCUS AX044034
DEFINITION Sequence 113 from Patent WO006791.
ACCESSION AX044034
VERSION AX044034.1 GI:11342918
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE
1 (bases 1 to 34980)
AUTHORS Piza,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C.,
Masignani,V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M.,
Scarlati,V., Rappoli,R., Frazer,C.M. and Grandi,G.
TITLE Neisseria genomic sequences and methods of their use
JOURNAL Patent: WO 006791-A 113 09-NOV-2000;
CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)

OY	1894	aaacgcttcttccagcgccagaccgcagccgacgcttaccatctttagaagcggtg	1953
Db	287277	AAACGCTTTTTCAGCGCGACACCAACCCGACCGCTTCAATTCATTTAAACGACCATTTGG	287218
OY	1954	tcaaaaaatgaagatctcccaacaagagaaatctgttggagcaaacagctatgtatcmaacc	2013
Db	287217	TTCGCAAAAAGAGGGCATTTCTCCGGGGGAATCTGTGTGGACACAGACTGGATCAACCGC	287158
OY	2014	acggttaaacgcggaaaattccatlatcaaggcgcggaacgcytgaattccgcgaatgtt	2073
Db	287157	ACATTTAAAGCGGAAAACTTCCAAATTTAAAGCGCGACAGCGGTGTTCGCCGAAATGTT	287098
OY	2074	ggcaaaagttagaagcgattncaattttagaataatcaagccccaagaagttttgttgtga	2133
Db	287097	GCCAAAGTGAAMAGCGATTGTGGCATTTAAGCATATACGCCCAAGCAGTTTGTGTGGGA	287038
OY	2134	ccgcatacaagaacatacaactctgtacacgcttcgagcttgaacngtctgaacaattgtctc	2193
Db	287037	CCGCATCAAAAGCCACACAACTGTGTACACGTTCCGAGCTGGACGGGTCTGCACAAATGTGTC	286978
OY	2194	gaanaaancattcccgagctaaagtgatgtctcatgtactgaacngactnaagcygc	2253
Db	286977	GAAMAAAACATTTCGCGACGATTAAGTATGCTTCATTGACTTAACACCGACATCAAGCGGC	286918
OY	2254	atgttaagcttmcnaatnagctnmttinaantcncnggcmgtmnaactnaangc	2313
Db	286917	AATGTGATCTTCCCATCAACGCTCATTTAAATCTCACAGGGCTTGCCACACTCAAGGCGC	286858
OY	2314	aacttagtgcgaatctgagctatcaagcttatcaagtcgaaccaacgcgaacccaacaaacygc	2373
Db	286857	AATCTTAGTGCAAATAGCGCGATACACGTTTATACATGCACGCCACAAAGCCACCAAAAACGCGC	286798
OY	2374	aacctaaagctctgtggcaatgcccgaagcaatttaacgaagcaacttaagcaacttaagcgcaac	2433
Db	286797	AACCTTAGCCTCGTGGGCAATGCCAAGGCCAACCAATTAACTTAACGCCAATTAACGGCAAC	286738
OY	2434	ncactgntctcgagcaatgtcttcaatttaacttaagcaacaacgcgcgacaacaaacygcagct	2493
Db	286737	ACATTCGGCTTCGGGCAATGCTTCATTTAATTAAGCGACACACGCCGTACAAAACGGCAGT	286678
OY	2494	ctgaagccttccgacaaacgcttaaggcaacgtaagccaattccgcactcaacgcycaatgtc	2553
Db	286677	CTGACCGTTTCCGGCAACGGCTAAGGCAAACCTPAAGCCATTCGCACTCAACGGTAATGTTC	286618
OY	2554	tcctctagccgataagcagtaattcattttgaacaaacgcgcgttttaacgcgaacaactcagc	2613
Db	286617	TTCCTTAGCCGATTAAGCAGATTTTCATTTTGAAGCAGCCGCTTTACCGGACCAAAATCAGC	286558
OY	2614	ggcagcaagaanaacagcatttaacttlaaagaacgcgcgaatvgacgctgcgctcaagcagc	2673
Db	286557	GCGCGCAAGCATACGCGCATTTACACTTTAAAGACAGCGAATGAGACCTCGCGTACAGCAGC	286498
OY	2674	gaattaggcaatttaaacctcttgacaacgcacccatlaacttaatttcgcctatgcgcgc	2733
Db	286497	GAATTAAGCAATTTAAACCTTGAACACGCCACCATTTACACTTAATTCGCGCTTATGCCAC	286438
OY	2734	gatgcttgcagcgcgcaaacaccgcgcagngtgcagacacgcgcgcgcgcgcttgcgcttgcg	2787
Db	286437	GATGCGCGACAGGGGCGCAAAACCGGCAGTGCAGACGATGTGGCGCGCGCGCTTGCGCGGT	286378
OY	2788	---cgccgttccctatatactcgcttaacacgcgcgaacttcgtgtgaatcccggttcaacag	2844
Db	286377	TGCGCGCGCTTCCCTATTATCCGTTACACCGGCACAACTTGGGTGAATCCCGCTTCAACACG	286318
OY	2845	ctgaacgctaaacgcgaatltgaacngtcaagaaacatccgcgttlatgtcgcgaaccttc	2904
Db	286317	CTGACGCTAAACGGCAATTTGACAGGTTGAGGGAACATTCGCTTTATATGTGCGGAACCTTC	286255
OY	2905	ggtctaccgaagcgacaacttgaagctgtgcgcgaaagtctccgaagmacttaacacttgcgc	2964
Db	286257	GGCTTACCGCAGCGCAAAATGCAATGAACCTGCGGCAAACTTCCGAAGGAGCACTTACACTTGGCGC	286198

[illegible]


```
Db 601 AGGCATATTTGGCGATTCGATGAAATGAGCCCAATTAACCGGAAAGTTTCATATCATATT 660
OY gca-----tggttaattggcgcaatacacatatalgcaggttgggaaatcat 699
Db 661 GCAACTGGGATTTCTGGCTGGTGGTGGCAATACCTTTCACAAATGATGACAGTGGT 720
OY ggcgtantatgtttgagcgcggtatgtgcgcatagcacaagactatggccctatgcgact 759
Db 721 GGCACAGTCATCTTATGATGTAATAAAATTAACATAGCCCTATATGATGTTTATACCA 780
OY gcaaggctgcgcagcgcaagcggttcgcgaatgttatltatltacgaacaaacaacataaa 819
Db 781 GGAGGCTCATTTGGCGACATGGTGGTCCACCAATGTTATCTATGATGATGCCCAAGCAAA 840
OY tggcgctcaacgaggttttacaacgcgtaccccttattcccggaaggaaacggttc 879
Db 841 TGGTTAATTAAATGGGATTTTGCACAAACGGCCATCTATATGGAATAAAGCAATGGCTTC 900
OY cagctgatacgcaaaagattgltctacgtatgcattacagagcgcatcacataccgttc 939
Db 901 CAGCTGGTCCGTAAGATTGGTGTATGATGAATCTTGTGAGATGCCCATTCAGTA 960
OY tttttgacgcgcgcagtaacgagacattttcttctacatccacaacagcggtacggt 999
Db 961 TTCTACGAACCGTCACAAATGGAATATCTTTTAAACGAGATATATATGACACAGA 1020
OY acggtatacgaagaacgaagaagunltccaaacaa---gctlaagtlacagacagtc 1056
Db 1021 AAAATCATATGCGCAACATGACACCAATTCCTGCTTAATGATTAATAAACGACGACGCT 1080
OY gcaagctttgacgaatcctttgaaacgaactgataaagaacagctt---acgcgcagag 1113
Db 1081 CATTTGTTTAATGTTTCTTTTATCCGACAGACAGAAACCGTGTATCATCTGACAGT 1140
OY ggtgttaatcagtlaccgttccaaaggttaaaacggtgaaaccttcttactatcgtatc 1173
Db 1141 GGTGTCAACAGTTATCGACCCGACTGAATTAATGGAATAATTTCTTTTATTAACGA 1200
OY ggcacgcgcgaacatccttcttcaacaacacacacacacgcgcgcgcgttcttctt 1233
Db 1201 GGAATAAGGGAATGTACTACTACGACCAATCATCAATCAAGGTGCGAGATTAATATTTC 1260
OY gaaagttatcttaacggttcgcgtcgaacaaacgaagaacgttgcgaagcgcggttctc 1293
Db 1234 CAAGGAGATTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
OY atcagtgaaagacagtlaccgttacttgaagaagtaaacgcgttgcgaacgcgcgttc 1353
Db 1321 ATCAGTGAAGACAGTACGTTACTTCTGGAATGAACGGCGTGGCAACGCCGCTGTC 1380
OY aaaaatcgcaaaagcgacgttcgacgttcaagcccaaaaggggaaacaaagctcgatcag 1413
Db 1381 AAAATCGGCAAAAGGCACGCTGACCTTCAAGCCAAAGGGGAAACCCAGGCTGATCAGC 1440
OY gtcgcgcagcgttaccgttcttctgaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1473
Db 1441 GGGGGGAGAGTACAGTCAATTTTGATGACGACGACGACGACGACGACGACGACGACG 1500
OY ttaagtgaatcgctgcttgcacgagcgaggttaacgttgcacgtacgtacgtacgtac 1533
Db 1501 TTTAGTGAATGCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1560
OY ttcacacccgcgaacatccttctcgcgttccgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1593
Db 1561 TTTCAACCCCGCAAACTCTATTTCGGCTTTCGGGGGAGCGATTGATTTAAACGGCGAT 1620
OY tgcgttcttctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1653
Db 1621 TCGCTTTCCTTCACCGATTTCAAAATACGATGAAGGGGCGCATGATTTCAACCAAT 1680
OY ggcacacacacacacacacacacacacacacacacacacacacacacacacacacacac 1713
Db 1681 CAAGCAAAAGATCCACCGTTTACATTAACAGGCAATTAAGATATTGTTACAAACCG --- 1736
OY aagaatcatatagacttaattatcacgaagaagaatltgcctacaacggttgcgcgcgc 1773
Db 1737 --CATTAACACACGCTTGATGATCAAAAAGAAATGCTTCAACAGCGTGTGGTTGGCGAG 1794
OY aagaatcacgaagaagaagaagcggttgcacacgttcttgcgcgcgcgcgcgcgcgcgcgc 1833
Db 1795 AAAGATACGACCAAAACACAGCGGCGCTCAACCTTGTATACACCGCGCGACAGAAAC 1854
OY cgcacccngctgtcttccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1893
Db 1855 CCGACCTCTGCTTCCGCGGGAACAAATTAACGGCAACATCCCAACCAACCAACCGC 1914
OY aaacgttlttccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1953
Db 1915 AAACGTGTTTTCAGCGGACAGACACACCGCGCTTCACATTCATTTAAAGACCATTTGG 1974
OY tcaaaatggaaggtatcccaagaagagaatcgttgcgcgcgcgcgcgcgcgcgcgcgcgc 2013
Db 1975 TCGCAAAAAGAGGCGATTCCTGCGGGGAAATCGTGGGACAAACGACTGATCAACCGC 2034
OY acgttcaagcgaaatcttccatatcagggcgagcggtgatttccgcgcgcgcgcgcgcgc 2073
Db 2035 ACATTTAAAGCGGAACCTCCAAATTAAGGCGGACAGCGCGTGTGGTTCCCGCAATGTT 2094
OY gccaagtggaaagcgatltmcatlttgagcaatcaacgcgcgcgcgcgcgcgcgcgcgcgc 2133
Db 2095 GCAAAAGTGAAGGCGGATTTGATTTGAGCAATCAGCCCAAGCGATTGTTGCTGCCA 2154
OY ccgcataaagcacaataccttctacacgcttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2193
Db 2155 CCGCATCAAGCGCACACATCTGTACAGTTGCGACGTGGAGCGGCTGTCAACAAATGTGTC 2214
OY gaanaancattaccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2253
Db 2215 GAAAAAACCTTACCGAGATTAAGTGAATGCTTATTTACCTTAACACCGCATCAGCGCG 2274
OY antgnaagntlmcnatlnaagntlmnaaancnnggntlmcnatlnaagntlmnaagnt 2313
Db 2275 AATGTGATCTTGGCGGATTCACCTTAAATCAACAGGCGCTTGCACACTCAACG 2334
OY aatctatgtcgaatltgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2373
Db 2334 AGCTTTAGGCAATGCGCATACGCTTATACGCTTACAGTACGACCTTACGCCAACCACAC 2394
OY aacctagcctcgttgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2433
Db 2395 AACCTTAGCCTCGTGGCAATGCCACAGCAACATTTAAATCAAGCCACATTAACCGCAC 2454
OY ncatcgntltgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2493
Db 2455 ACATGGGCTTGGGCAATGCTTCAATTTAAATGCAAGCGCGCTTAAACGACGAGT 2514
OY ctgaagcttccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2553
Db 2515 CTGACGCTTTCGGCAACGCTTAAGCAACGTAACGATTTCCGCACTCAACCGTATGTC 2574
OY tccctagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2613
Db 2575 TCCCTTAGCGGATTAAGCATTAATTTGAACGACGCGCTTAAACGACGAGTAAATATCAG 2634
OY ggcacaaagaaacacacacacacacacacacacacacacacacacacacacacacacacac 2673
Db 2635 GCGCGCAAGATACGCGCATTTACATTAATAAGACACGAAATGAGACGCTGCAATCAGCAG 2694
OY gaattaggaatttaaacctltgacaagcgacacacacacacacacacacacacacacacacac 2733
Db 2695 GAATTAAGCAATTTAACTTACACGCGACGACCATTTACCTCAATTTCCGCTATGCGCAC 2754
OY gatgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2784
Db 2755 AATGGCGACGCGGCGCAACCGCGAGTGGACAGATGCGCGCGCGCGCTTGGCGCGCT 2814
```


QY	1594	tcgcttcgttccacoglatccaataacogaygaagyggyatgatltgmcnatcataat	1653
Db	1621	TCACCTTCGTTCACACCTGATTCACAAATAACGATGAAGGGGCGATGATGTCAACCAAT	1680
QY	1654	gccaacaacatccacccgttaccattacgaagggatgaatgaatatacaacogaygt	1711
Db	1681	CAAGACAAAGAAATCCACCGCTTACCATTACAGGCATTAAGATATTGGTACAAACCG	1736
QY	1714	aagaatacaatagacttaatacagcaaaagaatltgcatacaogtltgtltbgcgag	1772
Db	1737	--CAATACAAACAGCTGTGATGACAAAAAAGAAATTGCTCAACAGGTGGTGGCGAG	1796
QY	1774	aaagatagagcaaaaacagaaacgggcygctcaacctgttttccagccgcgcagaaagac	1833
Db	1795	AAAGATACGACCAAAACGAAGGGCGGCTCAACCTGTTTAAACAGCCCGCTGCAGAAAC	1854
QY	1834	cgcaaccngctgtcttcgcgcggaacaattttaaocgcaacatacagcaacaagaagc	1896
Db	1855	CGCACCTGTGCTTCCGGCGGAGCAAAATTTAACGGCAACATCAGCAACAAACCGC	1911
QY	1894	aaactgttttcaagcgggcagaccgcaacgcgcatacaalttaagaaagcgggttg	1953
Db	1915	AAACTGTTTTACGCGGCAACCAACCCACCGCTACAAATCATTTAAACGACATGG	1974
QY	1954	tcaaaattgaaagttcccccaagagaaatcgttfgggacaacgcgtatcmaccgc	2013
Db	1975	TCCCAAAAAGAGGCAATTCCTCCGGGGAAATCGTGGGCAACAGCTGATCAACCGC	2034
QY	2014	acgtttaaagcggaaaaattccalatlacagggcgggaagcgggtatcccgcaatgt	2073
Db	2035	ACATTTAAAGCGGAAAACTTCCAAATTAAGGGCGAGAGGGGTTTCCGCGAATCTT	2094
QY	2074	ggcaaaagtgaagggcgaattgncatttgagcaatcaacgcaccaagcagttltgtgtcgca	2133
Db	2095	GCCAAAGTAAAGGCGCATTTGGCAATTTGAGCAATCACGCCCAAGCAGTTTTGGTGTGCA	2154
QY	2134	ccgcatacaagccatacaatctgtacaogcttgcagcggcngtctgcacaatltgtc	2193
Db	2155	CGGATCAAAAGCCACACAMTCTGTACAGCTTCGGAGCTGGAGGGGTCTGCACAAATTTGTCT	2214
QY	2194	gaanaaancattacgcagataaagtgtatgtcttcattgaactaagaacngacnlnagcgc	2253
Db	2215	GAATAAACCATTAACCCACGATTAAGATGATGTTTCATTTGACCAATACCGACATCGCGC	2274
QY	2254	atgtinaagctlnccaaatnagctlnltttaaactcncnggntlgmncactnaangc	2313
Db	2275	AATGTACGCTTGCCGATACGCTCATTTAAATCTCAAGCGGCTTGCCACATCTAACGGC	2334
QY	2314	aacttagtgcgaatlgcgatacaqgtlatacagtaagcacaacgacaccccaaaacgcg	2373
Db	2335	AATCTTAGTGCAAATGGCGATACAGCTTATACAGTACCCCAAGGCCACCAACCAACCGC	2394
QY	2374	aacctagcctgtlgygcaatgcgccaaagcaaatltaacagccacattaaacygcaac	2433
Db	2395	GACCTTAGGCTCTGTGGGCAATGCCAAGCAACATTTAATCAACCCACATTTAAAGCGCAAC	2454
QY	2434	ncatcgmgctgggcaatcttcaatttaactcnaagaaacagacgcgcgaacaaacgcgagt	2493
Db	2455	ACATGGCGCTTGGGCAATGCTTCATTTAATCTTAAGGACCAAGCCCGGCAACAAACGGCGAT	2514
QY	2494	ctgacgccttcgcgaacaacgctlaagcaaaqgtaagcatccgcactcgcacaaacygaatc	2553
Db	2515	CTGAGCGCTTTCGGCAACGCTTAAGGCAACGTAAGCATTTCCGCACATCAACAGGTAATGTC	2574
QY	2554	ttccctagccgaataaagcagtatltccaattltgaacaacgcgcgttaccggacaactcagc	2613
Db	2575	TCCTTAAGCGGATTAAGCGCATATTTCATTCGAAAGCAGCGCTTTACCGGCAAAATACAGC	2634
QY	2614	ggcaagcaaggaacagcaatatacacttaaaagacaggaatbgcgcgtgtgcgttaagcagc	2673
Db	2635	GGCAGCAAGGATACGGCATTTACACTTTAAAGACAGGCAATGGAGCGCTGACGTAAGGACG	2694

[illegible]

Db 1141 GGGGTCACAGTATTCAGCCAGACTGATATGAGAAAAATATTTCCTTATATGACAAA 1200
Qy 1174 ggcacgagcaactcatctatcaacaacatcaacaagcgcggttgatctt 1233
Db 1201 GGAAGAGGTAAATTTGATCTTACAGACATCAACCAAGCGGGGCTTTATTTT 1260
Qy 1234 gaagtgatttcaagtcgtcgtcgtcaacaacgaaacgttgcaagcgcggttcac 1293
Db 1261 GAGGATATTTTACGGCTGCTGCTTAAACACGAAACGTCGCAAGCGCGGCTTCAT 1320
Qy 1294 atgaatgaacagttaccgttacttgaaagtaaacgagcgtgaaacagcgcttc 1353
Db 1321 ATAGATATGCGATACCTTACTTGTGAAGTAAACGCGGGGAGACCGCTGTC 1380
Qy 1354 aaaaatcgcaaaagcagcgtcgtcgtcaagcgaaggggaaaaacgaagcgtcgtac 1413
Db 1381 AAATCGGCAAGGACAGCTGCTGCTTCAAGCCAAAGGGGAAACCAAGCTCGTCAGC 1440
Qy 1414 gtgggagcagtgatcatcttgatcagcagcagcagcagcagcagcagcagcagc 1473
Db 1441 GTGGGCGACGTAAGTATCTTATGATCAGCAGCGGACATCAAGCAAAAAACAGCC 1500
Qy 1474 tttagtgaatcgtgctgntcagcgagcggtgaagtgcaactgaatgcgataatcag 1533
Db 1501 TTTAGTGAATCGGCTTGTGACAGGCGAGGCGGCTGCACATGATGCGGATATCAG 1560
Qy 1534 ttaacccgcaaaactctatcttcgcttcggttcggtgagcagcgttgatcttaaacg 1593
Db 1561 TTTACCCCGCAACACTTATTCGCTTTCGCGGCGGAGCTTTGGATTGAAGCGGCAT 1620
Qy 1594 tgcgttcgttcacgcttctcaaaatacagatgaagggcgatgatgntcatcactac 1653
Db 1621 TCCCTTTCGTTACGCGCATTTCAAAATATCGATGAAGGGCGATGTTGCAACCAAT 1680
Qy 1654 gcaacaacacatccaccgttaccatcaatgaaggaatgaagatatacaaacgagcgt 1713
Db 1681 CAAGACAAAGATTCACCGTTACCATTTACAGGCAATTAAGATTAATCTTACACCGGC 1737
Qy 1714 aagaatatacagacttaatacagcaaaagaaatgctctcaaacggttggttcgag 1773
Db 1738 ---AATACCAACACTGATAGCAAAAAGAAATTCCTTCAACGCTTGGTGGGAG 1794
Qy 1774 aaagatacagcaaaagaaagcagcggttcaacctgttaccagcccgcgagaaagac 1833
Db 1795 AAAGATGCAACCAAAACGAGCGGGCTCAATCTGATTAACCAACCGGAAGCGGAT 1854
Qy 1834 cgcacccngctgttcccgacggaacaatlaaacgcaacatcagcaaacgaacgagc 1893
Db 1855 CGACCTTACTGCTTTCGCGGCAAAATTTAAACGCAATATACGCAACCAACAGCGC 1914
Qy 1894 aaactgttttccagcggaacgagcagcagcagcagcagcagcagcagcagcagc 1953
Db 1915 AAATGTTTTTACGGCGCAACCGCACCGCTTACATCAATATTAGAAAGCGGCTGG 1974
Qy 1954 tcaaaaaatgaagatcccaacaaggaatcgtgtggaacaacgagcagcagcagcagc 2013
Db 1975 TCAAAATGGAAGTATCCCAACAAAGGAATGCTGTGGGACAAACGATTTGATGACGCG 2034
Qy 2014 acgtttaaagcggaaatctccatctcagcgcgagcgagcggtgatctccgcaatgt 2073
Db 2035 ACATTTAAACGGAACAACTTCATATTCAAGGCGGACCAACCGGTGTTCCCGCAATGTT 2094
Qy 2074 gcaaaagtgaagcagatgncatttgacatcaacgccaagcagcttctgtgtcga 2133
Db 2095 GCCAAGTGAAGGCGATTTGATTAAGCAATCAAGCCCAAGAGTTCGCTGTCGCA 2154
Qy 2134 cgcacataaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2193
Db 2155 CCGCATCAAAAGCGCAACAACTTACACGTTGCGATGAGCGGGTTCGAAGATTTGAC 2214
Qy 2194 gaanaaanaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2253
Db 2215 GAAAAAACCATTTACCGACGATTAAGTATGCTTCAATTGAGCAAAAGACCGCATCAAGGCG 2274

Qy 2254 atgtinagctnnccnatacagntnnntnaaancntcngggcngcncactnaangc 2313
Db 2275 AATGTACGCTTGGCGATTCACGCTCATTTAAATCTCACAGSACTTGGCACATCAAGCGC 2334
Qy 2314 aatctagtcgaatlygcagatcacglttataagltcagcacaacgcccacaaacg 2373
Db 2335 AATCTTGTGACAGCGGAGACACGCACTATACGGTTACGGCAACGCGCCCAAAAGCGC 2394
Qy 2374 aactttagcctgtggcgaatctgcccgaagcaacatttaaacgacatlaaagcagc 2433
Db 2395 AACTTACGCTCGTGGCAATTCGCCAAGCAACATTTATATAGCGACATTAACGCGAAC 2454
Qy 2434 ncalcgnttcggcgaatlygtcatttaactaagcaacaacgcccacaaacgagc 2493
Db 2455 ACATCGGCTTTCGCACTATGCTTCTTAACTTAAGCAACAGCGCCGTACAAAGCGCAGT 2514
Qy 2494 ctgacgcttccgacaacgcttaagcaaacgtaagcagcagcagcagcagcagcagc 2553
Db 2515 CTGACGCTTTCGCACTATGCTTTCGCACTATGCTTTCGCACTATGCTTTCGCACTATGCT 2574
Qy 2554 tccctagcagataagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2613
Db 2575 TCCCTACCGCATTAAGCGCATTTCCATTTGAACACAGCGCTTACCGGAAAAATCAGC 2634
Qy 2614 ggcagcaagaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2673
Db 2635 GCGGCAAGGATACGGCATTAACCTTAAGCAACGCAATGAGCGCTGCGCGGCGCAGC 2694
Qy 2674 gaattagcaatttaaaccttgacaacgcaacatlaacatcaatlcgcctatcgcac 2733
Db 2695 GAATTAAGCAATTTAACTTGAACAGCGCACCATTTACCTCAATTCGCTTACGAC 2754
Qy 2734 gatgtcgaagcgcaaacgagcagcagcagcagcagcagcagcagcagcagcagcagc 2793
Db 2755 GATGCGGAGCGCGCAACCGGAGTGGCGCATGCGCGCGCGCGCGCGCGCGCGCGCG 2814
Qy 2794 tccctatlatcgttlaacacgcaactcgtgtagaalcgcgttcaaacagcgtgaagc 2853
Db 2815 TCCCTATATTCGTTACGCGCGCACTTGCAGAAATCCGTTCAACAGCGTCAGCGTA 2874
Qy 2854 aaagcgaatctgaacngtgaagaaacatctgcgttcttgcgaacgttcgcgttcc 2913
Db 2875 AAGGCAATTTGAACCGGTGAGGAACTTCGCTTATGTCGAGACTTTCGCTTACGCG 2934
Qy 2914 agcacaatctgaagcgtggaagcttcggaagaaatctacacttggtggtcaaat 2973
Db 2935 AGCGGCAATTTGAAGCTGGCGGAAAGTTCCGAAGGCACTTACACCTTGGCTCAACAT 2994
Qy 2974 accggaacgaacccgtaagccttcgataatgaacggttagtgaaggaagaaacaa 3033
Db 2995 ACCGGCAACGACCCGTAACTTCGAGCAATTCAGGTATGAGAGGAAAGCAACACA 3054
Qy 3034 cgcgtgtccgaaaacttaattcaacctgcaaaacgaacagcgtgagcgcgcgtg 3093
Db 3055 CCGGTGCGCAAAATCTTATTTACCTTCAACCAAGCAACAGCGATGCGCGGCGATGG 3114
Qy 3094 cgttacaactcatcgcgaagagcagcagcagcagcagcagcagcagcagcagcagc 3153
Db 3115 CGTTATGAGTTTATCCCAAGAGCGGAGTTCGCTGCTGATATTCGCTGCTGCTGCTGCT 3174
Qy 3154 gacgttccgacaacactcgcgaagcgagaa----- 3183
Db 3175 GAGTTCGCGCAAAACGTCGCGCAAGCGCGGAGAAACAGAGCCGCTTGAAGCAAAACAG 3234
Qy 3184 -----gcacaaacagcgagaaagaaagcaacgagcagcagcagcagcagcagc 3231
Db 3235 GCACAACTTGCAGCAACCAACAGAGCGGAAAGCAACGCGCAAAACCTTTGACGCGCTG 3294
Qy 3232 atgtgc 3291
Db 3295 ATTGGCGCGCGGCGCATGCGCACCGAAGAGCAAGAAAGTGTTCGCGCAACGCGCGCGCAG 3354


```
/db_xref="GI:1408569"
/translation="MKTVFRILNPLACISIGIVSOAGHTYFGIDYOYRDPFAENK
GKFTVGAONIKVYKKGOLVGTSMTKAMIDPEVSVSRNGVALAVNOXYIVSAHVNGY
TDVDFGAGNPDQHRFTYKIKRNNYKKNLHPEPDYHNRLKFTPEAPITMTS
NMNSSTYSDPTKPYBERRVIGSKGROPMRDNDQKQVACAYHILTGNHNOGAGNGY
SYLGDVRRKAGEIPLPLPAGSKGDSGSPMLTYDAKQWMLNGLIRENPEFGKENG
OLVKSFYDELFEEDHLNLSYLRAGNGVYITSGNDNGSGSLQKAGIYSEIKITLAM
SLPLEKDKVHNPRYDPNITSPRLNMGVETLFPMDKGSILFADIDOGAGLITLAME
NFTVSPNSOTWQAGIHSVENSITVKNVCEVHRDLKIGKGTILHVAQKNKISIS
VGQKVLLEOADOAGNKOAFSEIGIVSRECTYLONDKQDPTDKEFYRPRGRDLN
GHSITPRIONTDEGAMIVNNNTQOANVTTCGNSIVLPNGNNINKLIDYKREIYNG
WGETDKNKNHGRNLNLYKPTEDRTLLSGTNNKIGITQKGLIFSGREPAPHYN
HLKRWSEMEGIPQGEIYWDHWINRTEKAFNFOIKGSAAVSRVSSIEGMWTSNN
ANATGVVPMQOANTICRSDWTGLTQCKVLDPTKVINSLPTQIINSJINTDSTA
NVKGLAKNGNVLTLNHSOFTLSNNATQIGNIRLSDNATVADNANLNGNHLVADAO
FSLKNSHFHQIOGDKGTVLENAATMTMPSDTLLONLTNNSTITLNSAYSASNNT
PRRSLETETTPTSAEHRPNTILVNGKLSGGTQFQPSLSFGYKDKIKLSNDACDI
ILSVNRKREPETLEQTLVBSKNOPISDLKFTLLENDHDAGLRLKLVKNDGEFR
LHNPJKOEHLNDELVRAEQERLLEAKOVEPTAKTQTEPRVSRARAPEDPLPD
OSLLMLEAKOAELETAETOSKATKRVRSRAVSPDLDOSELEALAEVLIDAPQ
QSEKDRLOAEAEKOROKDLISRYNSNLSSELATVNSMLVSODELRLFLVDQOSA
VTNINADKRRYSDARAYOOKTNLROIYQKALANRIGAVSHSMDTPEDOVK
NHATLPMGSEAOYOMGDIQFGVNGTGISAKMAEORXTHKAIWYGNASIOFR
LQGLQIOPYEGVNRXYTERENYOSYEVRAKTPSLAFNRIANGIRDTYFTPLDNISVK
PYFRVTVVDSMANVOJTVMLVLOQPFGRYWKQEVGLKAEILHFOISAFISKSGSO
LGRQNNVGLGRW"
BASE COUNT      1496 a      776 c      890 g      1155 t
ORIGIN
Query Match      31.3% Score 1359.4; DB 1; Length 4317;
Best Local Similarity 59.0%; Pred. No. 3.4e-300;
Matches 2550; Conservative 0; Mismatches 1566; Indels 207; Gaps 8;
```

```
QY 673 acacatacgaggttgaggaaataatgagctanttagttgagcgcgatgtagccat 732
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 682 ACACACATACGCGGTGAGCAGGTATGATTCGTATTGTTGGAGCGGATGTTGGTAA 741
QY 733 gccaacactatgcccctatgcccatacgatgtagcgtagcgagcgaggtttcgcaatg 792
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 742 GCGGGAATAATGATCCATTACCATTCAGGCTCAAGGGGGACAGTGGTTCCTCGATG 801
QY 793 ttattatgacaaacaaacaaataatgtagtgcctcaacgtagtttacaacggtctac 852
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 802 TTTATTATATATGCTGAAAAAACAATAATGTTAATTAATGGATATTACGGGAAGCGAAC 861
QY 853 cctattccggcgaggaaaacggtttccagctgtagcgaagaagttagttcagatgac 912
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 862 CCTTTTGAAGGCAAGAAATGGGTTTCATATTGGTTGCGCAATCTTAATTT--TGATGAA 918
QY 913 attacagaggcgatlacacataccgctctntttgaacgcgcaatgacgacattttcc 972
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 919 ATTTTCGAAGAAGATTTACATFACATCACTTACACCCGAGCTGGTAATGAGTGCACGA 978
QY 973 ttacatccaaacaaacggttagcggtacggtacagaaacaaacgaagaagttccaat 1032
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 979 ATTAGTGAATGATATATGTCAGGGGCTTATACCTCAGAAATCAGGAATACCATACAGA 1038
QY 1033 ccaaagcttaagtlacagacagtcgcagctgttgcaagatcttgatgaaactgataaa 1092
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1039 ATTTAAATTTACGTTAGCAATATGAGTTT-----ACCTTTGAAAGACAGAGATTA 1089
QY 1093 gaaccagttacgcgagcgagggtgttaatcagtlaccgctcgaaggttaacaaacgttgaa 1152
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1090 GTTCATATATCTTGATATGACGAGACCTTAATATTATCTCCACGCTTAACCAATGAGAA 1149
QY 1153 aaccttcttatacgatlaacgcaacgcaactcatcttataacacatcaaccaa 1212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1150 ACCCTATATTATATGATGATCAAAAACAAGATCATTAATCTTCGATCTGCATTAACCAA 1209
QY 1213 ggcgcggcggttctgatttgaaggtgattttagcgttcgcgcctgcaaaaacgaagaa 1272
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1210 GGGCGCGGGGCTTATTATTATTGAGGGGTAATTTTACATATCTCCAAATTTCAACCAAACT 1269
QY 1273 tggcagagcgcggtgctcatatcagatgaaacagtlacggttacttggaaagttaacg 1332
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1270 TGGCAAGGAGCTGGCATATCATATGAGAAATAAGCACCGTACTTGGAAGATGTAATAAGGC 1329
QY 1333 gtggcaaacgacgacctgtccaaatccggaagaagcgctgacgttcaagccaaagg 1392
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1330 GTGGAACATGATCGACTTTCTTAAATTTGTTAAAGGAACATTGACGCTTCAAGCCAAAGGG 1389
QY 1393 gaaaaccaaagctcgatcagcgctgagcgtagcagtlacttltggaatcagcagcgagac 1452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1390 GAAATATAAAGGTTTCGATTCAGCGTAGCGGATGTAAGTCAATTTTGGAGCAGCGGACGAC 1449
QY 1453 gataaaggcaaaaaaacaacccctttagtgaatcgcgcttgnccaagcgaggttagcgtg 1512
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1450 GATCAAGGCAACAACAACCCCTTTAGTGAATTTGGCTTGGTTAGCGGCAAGAGGACAGTT 1509
QY 1513 caactgaatgcccataatcagtlcgaaccccgcaaacctcatttcgcttcgagcgga 1572
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1510 CAATTAAACGATGATTAACAATTTGATACCGAATAATTTTATTTTCGCTTTCGTGGTGGT 1569
QY 1573 cgttggatttaaacgggacatcgcttccgttccacggtatlcataataacgataagga 1632
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1570 CCGTTAGATCTTAAGGGGCAATCATTAACCTTTAAACGATATCCAAATATACGAGAGGGG 1629
QY 1633 ggcagatttgnatcalatagcccaacaacaaatccacggttaccatlaaggaatgaa 1692
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1630 GCAATGATTGTGAACCATATATCAACTCAAGCCGCTTAATGTCATCTTACTTGAGGAGAA 1689
QY 1693 agtattacaacacgagtgtagaataatcaatagacttaattacagcaagaagtggc 1752
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1690 AGCATTTGTTCTACCTAATGGAATATATATTATTAACCTTGATTACGAAAGAAATTTGCC 1749
```

[illegible]

Db	2650	CGTTTCAACBACATTGACAGTAAATGGTAAATTGATGGCCAAAGGCACATTTCCATTTACT	2739
Oy	2893	tcggaacctcttcggtctacccgaagcgacaaatctgaagcttgcggaaagtccgaaggnact	2952
Db	2740	TCATCTTATTTATTTGGCTTTAAAGACGATAAATTTAAATTTATCCAAAGACGCTAGGGCGAT	2799
Oy	2953	taaacctcttcgggtgcacaacatagccgaagaaacccgttaagctctgatacgaatttgcgta	3012
Db	2800	TCATATTATTATCTGTTCGCAACACAGGCCAAAGAACCCGAAACCCCTTGACGATTAATCTTGG	2859
Oy	3013	gtcgaaagggaaagaacaacaacgcgtctgcgaaaccttaattcaaccttcgcaaaacyga	3072
Db	2860	GTTTGAAGCAAAAGATATATCAACCGTTATATCGATTAACCTCAAAATTTACTTTAGAAATATAC	2919
Oy	3073	cagcttcgattgcggcgcggtgcgtttacaaactatccgaagaagcggaattccgcgtg	3132
Db	2920	CACGTTGATGCGAGGTGCATTAACGTTTAATTTAGTAAAGATGATGGCAATTTCCGCTTG	2979
Oy	3133	cataatccgylcaaaagaacaagagcttccgacaacatcgcgaaagc-----	3179
Db	2980	CATAAACCAATTAAGAGCGAGGAATTCACAAATGATTTAGTAAAGCAGACGACAAACAGAA	3039
Oy	3180	-----agaaagccaaanaaacagcggaanaaagaacaacgcgcaagaccttgcgctg	3231
Db	3040	CGAACATTATGAAAGCCCAACAAAGTTGAACCGACCTGCTTAACACACAAACAGGTAGGCCAAA	3099
Oy	3232	atgcgcgcggcgcgatgc---cgccaaagaacagaagcgttgcgaacgcgcgcg	3288
Db	3100	GTCGCGCTCAAGAAAGACGACGAGAGCGACTTTCCTGATACCTCGCTGATCAAGCCCTG	3159
Oy	3289	cngcgaaagcgcggaanaatctgcgcattatgcagcggaagagaagaanaaacgggtgcag	3348
Db	3160	TTAAACGATTATGAAAGCCCAACAAAGCTGAACCTACTGCTGAAACACAAAAAGTAAGCA	3219
Oy	3349	gcgataaagacagcgcnttgcgnaaacagcgcgaagcggaaacccgcgcgntacaaac	3408
Db	3220	AAAACAAAAGGCGCGGTCAAAAAGGACAGTCTTTTCGATCCCTCGTTGATCAAAAGC	3279
Oy	3409	gccttcccccgcgccgngcgccgcgcgcggaatttgcgcaacgcggaagcccaacgcgca	3468
Db	3280	CTGTTCCGATTATGAAAGCCGCACTTGAAGGTTATTTGATGCCCCACAGCAATCGGAAAAAGAT	3339
Oy	3469	ccctcaaccccaa-----ccgcagcgcgacctgtatnagccgttat	3507
Db	3340	CGCTACGCTCAAGAAAGACGGGAAAAACAACGCAAAACAAAAGACTGTATGACCGTTAT	3399
Oy	3508	gcgaatagcgtttgcgtgaatttccgcgaagctcaacagcgttctgcgcgtacaagac	3567
Db	3400	TCAAATATGCGGTTATGCAAGATTAATCGCAACAGTAAATAGTATGCTTCTGTTCACAAAT	3459
Oy	3568	gaattgcgcgcgttcttgcgaaagcgcgcgaacgngtlttgacaagcngcatcccg	3627
Db	3460	GAATTAGTACTCTCTTTTGTAGATCAAGACCAATCTGCCGTGTGGCAAAATATCCCAAG	3519
Oy	3628	nacaccaaacctccgttgcgaagatlttccgcgcctacgcgcaacaacacgcgccttgc	3687
Db	3520	GATAAAGACGCTATGATTTGATGCGCTTCGCTTATACAGAGAAAACGAACCTTACGT	3579
Oy	3688	caaatcgtatgcagaanaaacctcgcgacgcggcgcgctgcgcatacctgttlttcgcaaac	3747
Db	3580	CAAAATTTGGGTCAAAAAGCCTTAAGCTTAATGAGCAGATTTGGGGCAGTTTCTCGCATAGC	3639
Oy	3748	cggacccgaaacacatctcgaacgcgcgcgcgcgcgaactcgcgaacatcgcgcgcgttgc	3807
Db	3640	CGTTACGATATATACCTTTTGATGAACAGGTTAAATAACACGCGACATTAACGATGATGTCG	3699
Oy	3808	gttttcgggaatacgcgcatacgcgaagttgcgatactgcgatacgaagcgcgcggtt	3867
Db	3700	GGTTTGGCCAAATATCAATGGGGCGATTTTCAATTTGCTGTAACGCTGGGAACGGGAATC	3759
Oy	3868	agcagcgcgantctnctcagacgcgcatacgcgcaaaacccgcgcgcgcgtgtcatatc	3927

Db 3760 AGTCGAGTAAATGCTGAAGAACAAAGCCGAAATTCATCGAAAGGATAATTAT 3819
OY 3928 ggcattcagcagatataccgagcgcggttgcggagatcggacatcgacatcgatcgc 3987
Db 3820 GGGGTGATGCAAGTATACGATTCGCTTGAAGGCAATGGGCAATGGCCCTTATTTTGA 3879
OY 3988 gcaacgctatttcgcacaaagacgaltacccgctacgaaagacgaltacatcgccac 4047
Db 3880 GTTAATGCTATTATTATTTAGACGTAAGTAATTCATGAGTAAGTGAAGTGAACG 3939
OY 4048 cccggttccttcaacccgttaccggttgcggatgaagcagatcttcaatcaacgcg 4107
Db 3940 CCTAGCTTGCATTTATTCCTATTAATGCTGCGCTGAGTGAATTAACCTTATTCGCG 3999
OY 4108 ggcacacacatccac 4167
Db 4000 ACAGATATATATACGCTTAAAGCCTTATTCCTGCTCATATATGTTTCAAGCCT 4059
OY 4168 aaagtcgcaac 4227
Db 4060 AACGATCAAAACACACGTAATCTACGCTGCTGCAACACACATTTGACCTTATTTGCA 4119
OY 4228 ggcgaatcggttaaaagcgcgaatcaaaaggttcacgctgtccacacacacacac 4287
Db 4120 AAAGAAATGGGATTAAGGCAAAATTTTACATTTCCAAATTTCCGCTTTTATCTCAAA 4179
OY 4288 gccaagaagcgaacacggaacacacacacacacacacacacacacacacacacac 4347
Db 4180 TCTCAAGTTCACACACGCAACGCAACAAATGTTGGGCGTGAATTTGGCTATCGTTGG 4239
OY 4348 taa 4350
Db 4240 TAA 4242

RESULT 13
AR157320 AR157320 4319 bp DNA linear PAT 17-OCT-2001
LOCUS AR157320 Sequence 1 from patent US 6245337.
DEFINITION AR157320
ACCESSION AR157320
VERSION AR157320.1 GI:16218252
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4319)
AUTHORS St. Geme, J.W. III and Falkow, S.
TITLE Haemophilus adherence and penetration proteins
JOURNAL Patent: US 6245337-A 1 12-JUN-2001;
FEATURES
source 1.4319
BASE COUNT 1497 a 776 c 891 g 1155 t
ORIGIN

Query Match 31.0%; Score 1348; DB 6; Length 4319;
Best local Similarity 59.0%; Pred. No. 1.4e-297;
Matches 2551; Conservative 0; Mismatches 1565; Indels 210; Gaps 9;

OY 73 cctgctactatgacatagcctgctgctgctgctgctgctgctgctgctgctgctgctg 132
Db 81 CTTAATTTTAAACCGCTTCATTTAGGAGTATGATGCAAGCGTGGCGTGACAC 140
OY 133 acttatttcgcatcaactcaactcaactcaactcaactcaactcaactcaactcaactca 192
Db 141 ACTTATTTTGGATGTATTCACCAATATATATGCTATTTTCCCGAGATTAAGGAAATTC 200
OY 193 gcaagtcgagcgaagatcagatgagttacacacacacacacacacacacacacacacac 252
Db 201 ACAAGTGGGCTCAAAATATTAAGGTTTATTAACAAACAGGCGCAATTAAGTTGGCACTCA 260
OY 253 atgacaaagaccccgatgatgtatttctgctgctgctgctgctgctgctgctgctgctg 312

Db 261 ATGACAAAGCCCGATGATGATTTTCTGTATGTACAGTACAGCGCGTGGACGCTT 320
OY 313 gtcggcagatcaatatatttgaagcgttgacacataagcggcgtatacaacgttgattc 372
Db 321 GTTGAATATCAATATATATGAGCGTGGCAATACCTAGATATATACATGATTTT 380
OY 373 gtcggaagaaagaaatcccgatcagcagcgttttcttcaacaaatctgaaagaaat 432
Db 381 GGTGCAAGGGAACCAACCCCGATCAACATCGTTTACTTATTAAGTTTGAATTAAGCAAT 440
OY 433 aatataagcctgaacatcaccacctcaacagggcattancatagccggttgat 492
Db 441 AACTCAAAAGAAAT 500
OY 493 aaattgtcacagatgcagacacgttcgaaatgacagatgcagatgcagatgcagatgc 552
Db 501 AAATTCCTTACAGAACGCGCTCCATTTGATATGATCTTCGATATGATGATGATGATGAT 560
OY 553 tccgataaagaaataatcccgagcgtgtccgacatcggttcagacacacatctgct 612
Db 561 TCAGATATGAAACAAATATCCAGAACGTTCTGATCGCTGCGACGCGATTTTGGCGA 620
OY 613 tatgatgatcaac 672
Db 621 AATGATCAAGCAAGGCGAACAGTTCGCGTGCATATATATATATATATATATATATAT 680
OY 673 acacatatcagggcttgaggaaataatgagcgtatgcttgcgagcgtatgagcgtatg 732
Db 681 ACACACATATGCGCTGAGCAGGTAATGATGATTTCTGATTTGGAGGAGGAGTTCGTA 740
OY 733 gccaacgactatgcccctatgcgcatgacggtgcgacgacgacgacgacgacgacgac 792
Db 741 GCGGAGATATATGTCATTATACGATTTGACGCTCAAGGCGACAGTGGTTCGAT 800
OY 793 ttatattatgacaaacaaacaaataaagtctgctcaacagaggtttacaacccgctac 852
Db 801 TTTATTTATGATGCTGTAACAAACAAATGTTATTTATGATGATTTATGCGGAAGGCAAC 860
OY 853 ccttatccgcaaggaagaaacggttccagctgacgacaaagatctgcttcaacgagac 912
Db 861 CCTTTGCAAGCAAGGCAAAATATGCTTCAATGCTTCAATGCTTCAATGCTTCAATGCTT 917
OY 913 atttacaagagcgtatcaacatccgctcttcttgaacccgacgacgacgacatcttc 972
Db 918 ATTTTCAAAAGATTTAT 977
OY 973 ttacatccac 1032
Db 978 ATTATGTAAT 1037
OY 1033 ccaaaagcttaagatcacagacacgctcagcgttgcgaatccttgaatgaactgataa 1092
Db 1038 ATTAAATTTACGTTAGCAAAATATGATTTT-----ACCTTTGAAAGGAAAGATTA 1088
OY 1093 gaaccagtttacgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1152
Db 1089 GTTATATATCTAT 1148
OY 1153 aaccttcttlatcgaatcagcagcagcagcagcagcagcagcagcagcagcagcagc 1212
Db 1149 ACGCTATATTTATGATCAAAACAAAGGATATATATATATATATATATATATATATAT 1208
OY 1213 ggcggcggcgttcttgaatgagcgttgaatgagcgttgaatgagcgttgaatgagcgt 1272
Db 1209 GGGGCGGCTGCTTTATTTTGAAGGATATATATATATATATATATATATATATATAT 1268
OY 1273 tgcgaagcgcgagcgtcgtatcagtgagcagcagcagcagcagcagcagcagcagcagc 1332
Db 1269 TGGCAAGAGAGCTGCAAT 1328
OY 1333 gtcgcaaacgacgcgctgcacaaatcgcgaagagcagcgtgcagcgttcaagcgaaggg 1392

gene	/translation="MSVITHHNPKSRETLLENKGIOPILIELYLOKQSVNELQ SIARKGIDDDVROMKRTKDELYKSLNLDNLDSQALEPKAISSEHSLIERPIYINDK AKIRPEPTLEIL" /gene="H10237" /note="this region contains an authentic frame shift and is not the result of a sequencing artifact; similar to GB:00096 SP:P77406 PID:1621505 PID:1788838 percent identity: 55.75; identified by sequence similarity; putative" /pseudo complement(2003. .2980) /gene="H10239" complement(2003. .2980) /gene="H10239" /note="similar to SP:P19674 GB:X56175 PID:581231 GB:00096 PID:1773093 percent identity: 48.00; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="protein-export membrane protein (secf)" /protein_id="AAC21907.1" /db_xref="GI:1573204" /translation="MMKLTQDKDGHFIREINGIKLPPLTEPMKVKLGIIISALIM VISLEFITKGNMGIDETGVYEDTHDFSDIQRSLKHEGIESPIVOTGSVO DVMRLPASNNSTGEHYKSMLONDKDIQIRTEFGPNVGEELAQGVATLATL AMVLIVGSRFEMKRGFGSIASLADVIITLGFSAIQIEIDITFYVALISVGYSTN DSIVEDRVRENFRKIRLDITIDISLQSLSRITITSVTLVVMALFFFGPSI HNFSIALVIGIEFGYSSIFVAIAIAYDVGLRREHMPKVKDEIDELP" complement(2988. .4838) /gene="H10240" complement(2988. .4838) /note="similar to SP:P19673 GB:S68715 GB:X56175 PID:581230 PID:545175 percent identity: 59.74; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="protein-export membrane protein (secf)" /protein_id="AAC21908.1" /db_xref="GI:1573205" /translation="MLNRYPLMKNLVIFVAIGILSLPNYGEDPAVOIGTRGOE AMTSVGOVDVLTNNLPKTSIVLENGSLIARNTINDOLAKDKIAERIGNNYTTA LNLAPATPAMLSMFGANPMKMGIDLGVRFLMEVNNATLVKROLOJSLGRRLK EKIOTYAKNTEHFFGLITLANYSORAKAERIRIROLPTDIPEDDSIMLGSTAA LNEORPLAEONTTIRKRYAEIGVAEAVTORGAERIVIELEPGVDATARKELIGAT ATLERIVNONVTADAIISRMLEPADSEVKIDROGHFVALKRAYLGSEHTINSSGID OHSSTPOVSVTLTDSGEELMSQTKRYKRPMAVLVEYDNGKAYLSEKHTILKHEE VINAVTIGRFGSNFOITGVDSIAEAHNLSLTLKSGALIAPIQIWERKTIQPSLGAON VEQGINASLMGLVAIVAFMLFYKMEGVISFALVINIIVLVLGMSLIPGATLMPGI AGIVLIGMSVDANVLIERKEIRNGRSIOOAINGVGAFPSIFDANLTTLITAI ILXAVGTGPIQGFATLSIGVAISMFTAITGTRALVALVAGKQKLLLI" complement(4908. .5201) /gene="H10241" complement(4908. .5201) /note="similar to PID:599589 percent identity: 62.37; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAC21909.1" /db_xref="GI:1573206" /translation="MEOSPMSTLFIPIVIGLIFYPMIYRPOAKRKHKKILMSLAK GPEVLTAGVIGKTKVTBEGSDSIVIALNNTTETLINNMYIVSVLPKSLMSL" complement(5309. .5530) /gene="H10242" complement(5309. .5530) /note="similar to GB:I77117 SP:Q58397 PID:1591652 percent identity: 35.71; identified by sequence similarity; putative" /codon_start=1	gene	/transl_table=11 /product="conserved hypothetical protein" /protein_id="AAC21910.1" /db_xref="GI:1573207" /translation="MKIOLNLTALRCPIPLSARAKALNLDKNDLMLNLNESAVEN FSIFAENSVALVEQYVASEKEFIVILK" complement(5527. .6045) /gene="H10243" complement(5527. .6045) /note="hypothetical protein; identified by Genemark; putative" /codon_start=1 /transl_table=11 /product="H. influenzae predicted coding region H10243" /protein_id="AAC21913.1" /db_xref="GI:1573208" /translation="WQILEPQGFATWNEPIEMLYACHSKVKKRCRQLSLIPYLEKHG YTOAVLNDVEQILSYFNRAAPLHNDDELDLFPQVLKVAPOQTSTIDELKQHEYLKE NMNALSVOLEELISEQRODIDKHLIERFIQGYDRHLALEPLEFEMGRECLSDILTEM GKHMSARQVKE" complement(6110. .7258) /gene="H10244" complement(6110. .7258) /note="similar to GB:M63939 SP:P19675 GB:S68715 GB:X56175 PID:147966 percent identity: 81.60; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="trna-guanine transglycosylase (tgt)" /protein_id="AAC21911.1" /db_xref="GI:1573209" /translation="MKYELDKTSGNARRGRVLFERPOGTFVSYPFAPMPVGTGYKNG MPEVRATGAELIGNPFHMLRPGOEVMRKGDLDHDFQOMHRPLITDSGQFQVESL GKLKRTTEGCVKPONPINGERIPLSPKEMEIQDSDVIMTFDECTPYPAFEDYAK KSMESLNRMAKSRDRFDELKGNALFGLIIGGVFPELKRVSIGLVNIGDGYANG LAVGEPKEDMIRILEYICPQIPADKPRYLMGVGKPELDVGVGRGIDMPCVWPTRNA RNHGLFVTDGIVKIRNAKYRDIPTSPDPCDCYCKNRYKAYLHLDKCCELLGARLN TINLRYYQRLAEIRQAIEDDFDFVVEFYARMGKPVPPQLADKS" complement(7602. .8693) /gene="H10245" complement(7602. .8693) /gene="H10245"	gene	Query Match 28.9%; Score 1255.8; DB 1; Length 13818; Best Local Similarity 57.7%; Pred. No. 2e-276; Matches 2515; Conservative 0; Mismatches 1616; Indels 225; Gaps 10;	gene	Query 73 cctgcttaacttaagcatalgcctgctgcgtgcattctcccaagcttgaggcggaac 132 13610 CTAAATTTTAAACCGCTGTGTTTCATTAGGGATACATCACAAAGCCTGGGCAAGTCAT 13551	CDS	Query 133 acttatttgcgcatcaactccaatlaactcgaacttgcggaataaaggcaagctt 192 13550 ACTTATTTTGGGANTGCTACCAATATTATTCGTATTTTCCGAGAAATTAAGGGAATTC 13491	gene	Query 193 gcaatcgaggcggaagatattgaggtntacaacaaaggagggattgctgcggaatca 252 13430 ATGACAAACCCCGAGATGATTGTTTCCGTGTCGTACAGGCGGTGGCGGACATTA 13371	CDS	Query 253 atgacaaaggcccgatgatatttcttgctgctgcgttaacagggctggcgagcatg 312 13490 ACAATGCGGGCTAAATATTGAGCTTATTAACAAAGACGCAATTAGTTGGTCACATCA 13431	gene	Query 313 gtggcgcatcaatatattgtagcgctggcaatacagcgcgcttaacaacgcttatctt 372 13370 GTAGCGCATGATATATTGATGAGCGTGCACATTAACGGCGGATATAACGATGTTGATT 13311	CDS	Query 373 gtgctggaaggaagaaatcccgatccagcaccgcttctcttaaccaattgcaagaagaat 432 13310 GTGCAAGACGACGAACCCGATCAGACACGCTTACTTATCAAAATGTATAAAGAAAT 13251
------	---	------	--	------	---	------	--	-----	---	------	---	-----	---	------	--	-----	--

[illegible]

Accession	Gene	Protein	Accession	Gene	Protein
Db 3655	ltcgcgcctacccgcacacaacacgcacccctgcgcgaataatcgtatgacagaaaaactcgcgc	3714	QY 3655	ltcgcgcctacccgcacacaacacgcacccctgcgcgaataatcgtatgacagaaaaactcgcgc	3714
Db 10097	TTCCCGCGCTTATTCAGAGAAAGAAACGAACCTTCGTGCAAAATTTGGGGTGAACCACTTATGAT	10038	Db 10097	TTCCCGCGCTTATTCAGAGAAAGAAACGAACCTTCGTGCAAAATTTGGGGTGAACCACTTATGAT	10038
QY 3715	agcgagcgcgctgcgcatacccgcttctgcacacacgcggaccgcgaataacantcttcagacgacgc	3774	QY 3715	agcgagcgcgctgcgcatacccgcttctgcacacacgcggaccgcgaataacantcttcagacgacgc	3774
Db 10037	AATGGACGAATTTGGGGCGGTTTTCTCCGATAGCCGTTTCAGTAATATACCTTTGACGAACAG	9978	Db 10037	AATGGACGAATTTGGGGCGGTTTTCTCCGATAGCCGTTTCAGTAATATACCTTTGACGAACAG	9978
QY 3775	atcgcgaactcgcgcacgcgctctgcgcacgcgcgcgcttctgcgcaatacgcgcatacgcgcgcgcg	3834	QY 3775	atcgcgaactcgcgcacgcgctctgcgcacgcgcgcgcttctgcgcaatacgcgcatacgcgcgcgcg	3834
Db 9977	GTTTAAATAATCAGCCGACGATTAACGATGATGTCGGGTTTTGTGCCAATATCAATGAGGGCGCAT	9918	Db 9977	GTTTAAATAATCAGCCGACGATTAACGATGATGTCGGGTTTTGTGCCAATATCAATGAGGGCGCAT	9918
QY 3835	ctcgcacatcgcgcatacgc	3894	QY 3835	ctcgcacatcgcgcatacgc	3894
Db 9917	TTTAACTTTTGGTGTAAATGTGGGTGGCGGGAAATTAATGTCGAGTAATAATGGCTTAAGAACAA	9958	Db 9917	TTTAACTTTTGGTGTAAATGTGGGTGGCGGGAAATTAATGTCGAGTAATAATGGCTTAAGAACAA	9958
QY 3895	gagagcaaaaatcgc	3954	QY 3895	gagagcaaaaatcgc	3954
Db 9857	AGCGCAAAATTCATCGAATAAGCGATAAATATGAGATGAATGCAAGTTATCATGTTCCGT	9798	Db 9857	AGCGCAAAATTCATCGAATAAGCGATAAATATGAGATGAATGCAAGTTATCATGTTCCGT	9798
QY 3955	ctcgc	4014	QY 3955	ctcgc	4014
Db 9797	TTTGGGCAATTTGGGTATTCACCCCTATTATTGGGTGTAAATGATTAATTTATGTAACCTGAA	9738	Db 9797	TTTGGGCAATTTGGGTATTCACCCCTATTATTGGGTGTAAATGATTAATTTATGTAACCTGAA	9738
QY 4015	gataccgc	4074	QY 4015	gataccgc	4074
Db 9737	AATTTATCAATTCGAGAAAGTAAAGTGAACACCGAGCCTTGCAATTAATCGCTATTAAT	9678	Db 9737	AATTTATCAATTCGAGAAAGTAAAGTGAACACCGAGCCTTGCAATTAATCGCTATTAAT	9678
QY 4075	gcgcgcataaagcagataltatcattcaaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	4134	QY 4075	gcgcgcataaagcagataltatcattcaaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	4134
Db 9677	GCTGGCGATTGAGTGTATTAATCAATTAATCCCGACAAATAAATATCAAGCTTAAT	9618	Db 9677	GCTGGCGATTGAGTGTATTAATCAATTAATCCCGACAAATAAATATCAAGCTTAAT	9618
QY 4135	tttagcctgctccttacccgagtcgc	4194	QY 4135	tttagcctgctccttacccgagtcgc	4194
Db 9617	TTCTTTGTCAATTAATGTTGATGTTTCAAAACGCTAACGTAACAACTGTAAATATGACAGC	9558	Db 9617	TTCTTTGTCAATTAATGTTGATGTTTCAAAACGCTAACGTAACAACTGTAAATATGACAGC	9558
QY 4195	glatgctcgc	4254	QY 4195	glatgctcgc	4254
Db 9557	ATGTTGCAACAATCATTTGGCGGTTATTGGCAAAAAGAGTGGATTAAGCAAGAAATT	9498	Db 9557	ATGTTGCAACAATCATTTGGCGGTTATTGGCAAAAAGAGTGGATTAAGCAAGAAATT	9498
QY 4255	aaaggttcaacgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	4314	QY 4255	aaaggttcaacgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc	4314
Db 9497	TTTAACTTTCCAACTTTCGCTTTTATTCTTAAATCTCAAGGTTCGACACTGGCAAAACAG	9458	Db 9497	TTTAACTTTCCAACTTTCGCTTTTATTCTTAAATCTCAAGGTTCGACACTGGCAAAACAG	9458
QY 4315	cacgc	4350	QY 4315	cacgc	4350
Db 9437	CAAAATGTGGCGTGAATAATGGGATATCGTTGTA	9402	Db 9437	CAAAATGTGGCGTGAATAATGGGATATCGTTGTA	9402

Db 421 CGTTATGATGATGACGACAAACAGCGCATTTATCTCTACTCCGCGCATGTTAATTGGC 480
Qy 667 ggaataacacatacagaggttggggaataatgagcgtantagttgagcgagatgtg 726
Db 481 GGCAATACACATATGACAGGTTGGGGAATAATGCGCTAGTTAGTTGACGCGCATGTG 540
Qy 727 ggcactgcgaacgactatgagcctatgcgagatgaggtgagcg 770
Db 541 CGCCATGCCAACGACTATGCGCTATGCGGATTCAGAGTGGCG 584

RESULT 17
NLGA823 572 bp DNA linear BCT 21-MAY-1998
LOCUS Nisseria lactamica Iga protease homologue gene, isolate NL823.
DEFINITION AJ001737
ACCESSION AJ001737.1 GI:3152389
VERSION Iga protease.
KEYWORDS Nisseria lactamica.
SOURCE Nisseria lactamica.
ORGANISM Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 572)
AUTHORS Sayers, J.R.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1997) Sayers, J.R., Division of Molecular
Medicine, University of Sheffield, Royal Hallamshire Hospital,
Sheffield, S. Yorks, S10 2UF, ENGLAND
2 (bases 1 to 572)
AUTHORS Sayers, J.R., Vltowski, S. and Read, R.
TITLE Strains of Nisseria lactamica carry an Iga1 protease gene
homologue and some express Iga1 protease activity
JOURNAL Unpublished
FEATURES
source Location/Qualifiers
1..572
/organism="Nisseria lactamica"
/isolate="NL823"
/db_xref="taxon:486"
CDS
1..304
/codon_start=2
/transl_table=11
/product="Iga protease homologue"
/protein_id="CAA04962.1"
/db_xref="GI:3152390"
/db_xref="SPTREMBL:O69748"
/translation="KDI EYNNKKGELVGKSMTRAPMTDFSVSRNGYALVGDQYIVS
VAHNGYNNVDFGAEKGNPDQHRFSTQIVAKRNKTPDNSHPYNGDHPRLPGISS"
repeat_region 44..51
/rpt_type=INVERTED
repeat_region 68..75
BASE COUNT 158 a 124 c 157 g 133 t
ORIGIN

Query Match 11.9%; Score 517.6; DB 1; Length 572;
Best Local Similarity 95.1%; Pred. No. 1.2e-107;
Matches 544; Conservative 0; Mismatches 23; Indels 5; Gaps 1;

Qy 204 gaaagatatgaggttacaacaacaaagggaggttggcggcaaatcaatgacaaaagc 263
Db 1 GAAAGATATTGAGGTTTACAAACAAAAGGGAGTTGGTCGCAATCATGACAAAAGC 60
Qy 264 ccgagatgatatttcttcgtggttcgctgaagcggtgctgagcattgtgtggcgatca 323
Db 61 CCCGATGATGATTTTCCTGCGTGTGCGGTAACGCGCTGATTAACAAATGTTGATTTGCGGAGCG 180
Qy 324 atatatgtgagcgltgacacaaacgagcgtatatacaacagttgatttgytgcggaag 383
Db 121 ATATATGTGAGCGTGGCGCATACGCGCGCTATTAACAATGTTGATTTGCGGAGCG 180
Qy 384 aagaaatcccgatcagcagcgttttcttaccgaattgtgaaaaaataatataagcc 443
Db 181 AGGCAATCCGATCAGCACCGCTTTTACCAAAATTTGTAAGAAATAACTACCAAGC 240

Qy 444 tgacaattcacaccccttaacacgagcattancatattgcccgcg-----tttgataaat 498
Db 241 TGACATTTACACACCCCTTTACACGCGCATTTACATATGACCGCGCTCCGTGCAATTAAGTTTC 300
Qy 499 gtcaacagatgcagaaccclytgcgaatgacagatgacatgagggggaataactatccgat 558
Db 301 GTAACCGATGACGAACCTGTCGAAATGACGGGAGACATGAGGGGGAATACCTATTCCGAT 360
Qy 559 aaagaaaaatccccgagcggtgtccgcatcgctccagacacacacatttggcgttatgat 618
Db 361 AAAGAAAATATCCCGAGCGGTGTCGCATGCGTCAGACACCACTATTGGCGTTATGAT 420
Qy 619 gatgacaacacgagcattatctactccgcgcgcatgttatttgcgcgcaataacat 678
Db 421 GATGACAAACAGCGCGCATTTATCTACTCCGCGCATGTTGATTGGCGGCAATACAT 480
Qy 679 atgcagaggttggggaataatgagcgtantagtttgcagcgagatgtgcgcatgccaac 738
Db 481 ATGCAGGCTTGGGGAATATGAGCGCTAGTTGAGGCGGATGTCGCCATGCCAAC 540
Qy 739 gactatgcccctatgcccattgcaagtgagcg 770
Db 541 GACTATGGCCCTATGCGCATTCGAGTGGCGG 572

RESULT 18
NLGA3293 572 bp DNA linear BCT 21-MAY-1998
LOCUS Nisseria lactamica Iga protease homologue gene, isolate NL3293.
DEFINITION AJ001738
ACCESSION AJ001738.1 GI:3152385
VERSION Iga protease.
KEYWORDS Nisseria lactamica.
SOURCE Nisseria lactamica.
ORGANISM Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 572)
AUTHORS Sayers, J.R.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-1997) Sayers, J.R., Division of Molecular
Medicine, University of Sheffield, Royal Hallamshire Hospital,
Sheffield, S. Yorks, S10 2UF, ENGLAND
2 (bases 1 to 572)
AUTHORS Sayers, J.R., Vltowski, S. and Read, R.
TITLE Strains of Nisseria lactamica carry an Iga1 protease gene
homologue and some express Iga1 protease activity
JOURNAL Unpublished
FEATURES
source Location/Qualifiers
1..572
/organism="Nisseria lactamica"
/isolate="NL3293"
/db_xref="taxon:486"
CDS
1..570
/codon_start=1
/transl_table=11
/product="Iga1 protease"
/protein_id="CAA04963.1"
/db_xref="GI:3152386"
/db_xref="SPTREMBL:O69745"
/translation="KDI EYNNKKGELVGKSMTRAPMTDFSVSRNGYALVGDQYIVS
VAHNGYNNVDFGAEKGNPDQHRFSTQIVAKRNKTPDNSHPYNGDHPRLPGISS"
repeat_region 67..74
/rpt_type=INVERTED
repeat_region 118..127
BASE COUNT 169 a 119 c 154 g 130 t
ORIGIN

Query Match 7.7%; Score 332.8; DB 1; Length 572;
Best Local Similarity 73.6%; Pred. No. 2.3e-65;


```

/codon_start=3
/transl_table=11
/product="Iga1 protease precursor"
/protein_id="AAC45792.2"
/db_xref="GI:12597969"
/translation="LAVALTPYSEALVVDVYDIOIFRDEAKKGFVGAUDLSVK
KOGONIGNALSNVPMIDSVADVNRRTLTVIDPOYAVSKHVKGEISYGHNGHLD
VENDEREVSVAONDEYEPKNNHNGNOGLEDYNNARLKKPTEVAPRIPASGCGVE
TYKDKRSESEPVACAGTOPEKNSRYNMELELSRATYAIAGPIQDVANTSLNDECL
IFGDNKSHSPKLEKLEYSQALNTYAVLGSGLFVYDQEKWVGLAVDWAG
YKNSMOEWNTIKKEPADIKORDNAGTICKGEHWNITFETNSKIGSTAVRLNGE
KANNQONVTEFEDNGLVDONINOGAGGLFPERGDTYVGINNDITMGLADVDGK
KYVQVKNPNDRRLAKIKGTELEINGVNOQLKVGDTVLINQADADKRVQAFSQ
VIGSGRGTVLNNSNOINPDMLFEFGRGRLDANNDLTFEHIRVDGCAIVNHT
SHASTITLTGKSLITNPNSLSYHSIONDEDDYSYRRPRPIPOGRDLXYKNRY
ALSGSVNAPMPENGVTENNVMVEGTQOEAKKAMKNNKRNORTSGSGFGEENG
KGNALNINFGKSAONRELLTGCTNLNGKLSYNOGNTLISCRPHARDVKNSSA
KRDARHSKNEVEFEDDMINRTFKAEIAYVNSASFSRSSRNSNTTATDNKAVN
LYKNGDEVASDITGTYCTNGLNSDKALNSFCATQINGVNLNQNALVLGKAL
WQIOQGNRSVSLNOSHSMWLTGDSQVNLNLSLASHILNNAQSNQANQYHLLKIN
HLSGNCHFYTLHLAENLDKVLKESASGHQLVHODKDTGEPNGEGLNLPDASVOD
RSRLSYSLANNHVDLGLRITIKTEGNTIRLYNPAENRRRVPAPSPATNTASQAK
ATOTDGAOTAKPQNTIVVAPPSPQAN"
BASE COUNT      984 a      683 c      642 g      686 t
ORIGIN

```

```

Query Match      4.5%; Score 197.6; DB 1; Length 2995;
Best Local Similarity 61.0%; Pred. No. 2.4e-34;
Matches 336; Conservative 0; Mismatches 212; Indels 3; Gaps 1;

QY 1161 ttatcatgatacgaacgaacactcatctatcaacaacacacgaacgaacgagc 1220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1028 TGTATCTTTGAAGACACACGCGCTTTGGTATTGATCAAAACATCAGCAACGCGG 1087
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1221 cggtttgatattgaagtgatattacggtctcgctgaaacaacga---aacgtggca 1277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1088 CGGTCTGTTTTTCAAGGCGCATTTACACAGTCAAGGATTAATTAATGACATCTTGCGT 1147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1278 aggcgagcgatcatcatcagtaagaacgtaaccgtacttgtaaaagtaaacgagtc 1337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1148 AGGTGCGGGGATTGATGTTACGACGCGCAAAAAGTCGTTGGCAGTCAAAATCCGGA 1207
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1338 aaacgacgcctgtccaaatccgcaaacgacgctgcgaacgtccaaaggggaaaa 1397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1208 TGGCGACGATTTGGCAAAATTCGCAAAAGGACACATTGAAATTAACGCGCACGCGCTTAA 1267
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1398 ccaagcctcgatcagcgtgagcaggtacagtcattttgagtcagcagcagcagcagataa 1457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1268 CCNAGGCGCATTTAAAGTTCGCGACGCTACGGTTATTCTGTAATCAACAAGCCGATGCCGA 1327
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1458 aggcgaacaaacgaacgtcttagtgaatcggtctgntccagcagcaggggtacggtcgaact 1517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1328 CAATAAAGTCCAGGCTTTCTCCCAAGTCGGCATTTGTCAGCGGACGCGGATCATTTGGTATT 1387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1518 gaatgcgatacagttcaaccccgcaaacctatcttgcgtcttcgagcgagcagttt 1577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1388 AATAGTCAATCAATGATTAATCCCGATACTATATTTTCGCTTCGCGCGCGTGGTTT 1447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1578 ggaattaaacggcagtcgcttcgttccacggtatccaataacgaatgaagggcgat 1637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1448 GGAATGCCAATGCAATGACTTTGAAACACATCCGCAACGTAAGTGAAGGCGCGG 1507
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1638 gattgcnatcataatgcacaacacatccacggttacatcatcaggaatgaagat 1697
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1508 CATGTTCACACCAACAAGCCAGCTCCACATACACGTTAAGCGGTAAATCTTTAAT 1567
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1698 tacacaacga 1708
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1568 TACCAATCCGA 1578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 22

```

AF235032
LOCUS      AF235032                4659 bp    DNA        linear    BCT 02-MAR-2001
DEFINITION Nisseria meningitidis Iga1 protease gene, complete cds.
ACCESSION  AF235032
VERSION     AF235032.1
KEYWORDS   GI:13183004
SOURCE      Nisseria meningitidis.
            Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria meningitidis.
FEATURES
    source   1..4659
              location/Qualifiers
    CDS       1..4659
              /codon_start=1
              /transl_table=11
              /product="Iga1 protease"
              /protein_id="AAK15023.1"
              /db_xref="GI:13183005"
              /translation="MKTRFRKFINAISIPLAVALTPYSEALVVDVYDIOIFRDEAK
              NKGFVATDLSVKKOGONIGNALSNVPMIDSVADVNRRTLTVIDPOYAVSKHVKGEISYGHNGHLD
              VENDORVSVONDEYEPKNNHNGNOGLEDYNNARLKKPTEVAPRIPASGCGVE
              TYKDKRSESEPVACAGTOPEKNSRYNMELELSRATYAIAGPIQDVANTSLNDECL
              IFGDNKSHSPKLEKLEYSQALNTYAVLGSGLFVYDQEKWVGLAVDWAG
              YKNSMOEWNTIKKEPADIKORDNAGTICKGEHWNITFETNSKIGSTAVRLNGE
              KANNQONVTEFEDNGLVDONINOGAGGLFPERGDTYVGINNDITMGLADVDGK
              KYVQVKNPNDRRLAKIKGTELEINGVNOQLKVGDTVLINQADADKRVQAFSQ
              VIGSGRGTVLNNSNOINPDMLFEFGRGRLDANNDLTFEHIRVDGCAIVNHT
              SHASTITLTGKSLITNPNSLSYHSIONDEDDYSYRRPRPIPOGRDLXYKNRY
              ALSGSVNAPMPENGVTENNVMVEGTQOEAKKAMKNNKRNORTSGSGFGEENG
              KGNALNINFGKSAONRELLTGCTNLNGKLSYNOGNTLISCRPHARDVKNSSA
              KRDARHSKNEVEFEDDMINRTFKAEIAYVNSASFSRSSRNSNTTATDNKAVN
              LYKNGDEVASDITGTYCTNGLNSDKALNSFCATQINGVNLNQNALVLGKAL
              WQIOQGNRSVSLNOSHSMWLTGDSQVNLNLSLASHILNNAQSNQANQYHLLKIN
              HLSGNCHFYTLHLAENLDKVLKESASGHQLVHODKDTGEPNGEGLNLPDASVOD
              RSRLSYSLANNHVDLGLRITIKTEGNTIRLYNPAENRRRVPAPSPATNTASQAK
              ATOTDGAOTAKPQNTIVVAPPSPQAN"

```

```

BASE COUNT      1542 a      1074 c      1041 g      1002 t
ORIGIN
Query Match      4.5%; Score 197.6; DB 1; Length 4659;
Best Local Similarity 61.0%; Pred. No. 2.6e-34;
Matches 336; Conservative 0; Mismatches 212; Indels 3; Gaps 1;

QY 1161 ttatcatgatacgaacgaacactcatctatcaacaacacacgaacgaacgagc 1220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1074 TGTATCTTTGAAGACACACGCGCTTTGGTATTGATCAAAACATCAGCAACGCGG 1133
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1221 cggtttgatattgaagtgatattacggtctcgctgaaacaacga---aacgtggca 1277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1134 CGGTCTGTTTTTCAAGGCGCATTTACACAGTCAAGGATTAATTAATGACATCTTGCGT 1193
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


RESULT 24									
LOCUS	NMG1471	1489 bp	DNA	linear	BCT 14-MAR-1995				
DEFINITION	N meningitidis iga gene (EM3771).								
ACCESSION	X82468								
VERSION	X82468.1 GI:732865								
KEYWORDS	iga gene; Iga1 protease.								
SOURCE	Neisseria meningitidis.								
ORGANISM	Neisseria meningitidis								
	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;								
REFERENCE	1 (bases 1 to 1489)								
AUTHORS	Lomholt,H., Poulsen,K. and Kilian,M.								
TITLE	Comparative characterization of the iga gene encoding Iga1 protease in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus influenzae								
JOURNAL	Mol. Microbiol. 15 (3), 495-506 (1995)								
MEDLINE	95302961								
REFERENCE	2 (bases 1 to 1489)								
AUTHORS	Lomholt,H.								
TITLE	Direct Submission								
JOURNAL	Submitted (19-OCT-1994) H. Lomholt, Dept. of Medical Microbiology, The Bartholin Building, University of Aarhus, 8000 C Aarhus C, DENMARK								
FEATURES	Location/Qualifiers								
source	1..1489								
	/organism="Neisseria meningitidis"								
	/strain="EM3771"								
	/db_xref="taxon:487"								
gene	1..1489								
	/gene="iga"								
	<1..>1489								
CDS	/gene="iga"								
	/codon_start=1								
	/transl_table=11								
	/product="Iga1 protease"								
	/protein_id="CAA57851.1"								
	/db_xref="GI:732866"								
	/db_xref="SPREMBL:Q57309"								
	/translation="ATDLVKNKOGKGINALSNVPIDESVDVNPRLTLVIDPQYA VSVHVKGDEISYGHNGHLDVSDNENERSVONQDEPKNMHNGORLEEDYNMA RLNFEYVADNPVIAPTSGAGGVEYKDKNRSEPRVAGQOFENSRVNMTELSRAVR YALNGTPYOVNNTSNLNOGLIGFGNSHHEPEKLEVISOLALTVYVIGDSSP LEAFDKEKRWFLGAYDVNAGYKKNQSMOWMNTYKKEFADRIKRDNAGLTKGGEH WKTGTQSHIGSTVRYRLAGNDRKNNKQNTFEDNGLTVLDQNIQSGGLEFGDITV VKGNNDITWLGAGIDVADGKRVVQYKNPMDGLATIGKTEINSGVNOQLKVG DGVLTWQADADKRVQAFSQVIGVSGRGLVLVNSMQINPDNLTFEFGRGRLDNGN DLTFEHLRVNDEGARIVNHNMTDRASTITLTKGSLITAPOMLSVEI"								
BASE COUNT	492 a 321 c 325 g 351 t								
ORIGIN									
Query Match	4.4%: Score 192.8; DB: 1; Length 1489;								
Best Local Similarity	60.8%: Pred. No. 2.8e-33;								
Matches 330;	Conservative 0; Mismatches 210; Indels 3; Gaps 1;								
Y 1161	ttttatcgattacggcaacgcaaacatcttatcaacaacatcaacgaagcgcg 1220								
Db 915	tggtaccttttagacgaacgacacttttggttgatcaaaacatcaacgaagcgcg 974								
Y 1221	cggattgatattgaagtgatttaagagtcgcgcgtgaagaacaacga--aacgtgca 1277								
Db 975	cggctgtgttttcaaaagcgcatctacacagtcgcaaaagtcgaatattatgcatcattggcct 1034								
Y 1278	aggcgcgagcgcttcataatcaatgaagacagtcacgttacttggaagtaaacgcgctggc 1337								
Db 1035	aggttcgggggatttgattgtttggcgcagcgcaaaaagtcgtttggcgaagtcgcaaaatccgaa 1094								
Y 1338	aaacgacgcctgctcaaaaatcggcaaaagcgacgctgacgttcaagcgaaggggaa 1397								
Db 1095	tgggcacaagatttggcaaaaatcggcaaaagcgacatattaaataataaacggcacagggctttaa 1154								
Y 1398	ccaaggtcgcatacgctgtggcgacgctgatacgtatttgatcaagcagcgacagaataa 1457								

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	source
Db	1155	CCAAGGCGCATTTAAAGTCGGGACGGTAGCGTTATTCTGAATCAACAAGCCGATGCCGA	1214										
Oy	1458	aggcaaaaaaacaagccttagtgaatcggcttgnctcagcgaggglaacgltgcaact	1517										
Db	1215	CAAAAAAGTCCAGCGTTTCTCCAGATGCGGCAATTCACGCGACCGGTACATTTGGTATT	1274										
Oy	1518	gaatgcgcatatcagttcaaccccgacaacatcatttcgcttcgcgcgagacgtt	1577										
Db	1275	AAATAGTTCACAAATCAGATTAAATCCCGATTAATATTTCGGTTCCGTGGCGGTCGTTT	1334										
Oy	1578	ggatttaaacggcgcatcgcttcgltccacccgatttcaaaatcacgatgaagggcgat	1637										
Db	1335	GGATTCACCATGGCATGACTTGACTTTTGACACATCCGACATCGATGAAGGCGCGCG	1394										
Oy	1638	gatttcacatcattatgcacacaaacatccacggttcacattacaaggaaatgaagat	1697										
Db	1395	CATTTCACACACACACACAGACCGCGCTCCACATATCATTACCGGCAATTCATTAT	1454										
Oy	1698	tac 1700											
Db	1455	TAC 1457											
RESULT 25													
NMIGAFA8													
LOCUS	NMIGAFA8	1489 bp	DNA	linear	BCT 14-MAR-1995								
DEFINITION	N.meningitidis iga gene (HF48).												
ACCESSION	X82480												
VERSION	X82480.1	GI:732875											
KEYWORDS	iga gene; 1941 protease.												
SOURCE	Neisseria meningitidis.												
ORGANISM	Neisseria meningitidis												
REFERENCE	Neisseria.												
AUTHORS	1 (bases 1 to 1489)												
TITLE	Lomholt,H., Poulsen,K. and Kilian,M.												
JOURNAL	Comparative characterization of the iga gene encoding 1941 protease												
REFERENCE	in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus												
AUTHORS	influenzae												
JOURNAL	Mol. Microbiol. 15 (3), 495-506 (1995)												
FEATURES	2 (bases 1 to 1489)												
source	Lomholt,H.												
gene	Direct Submission												
CDS	Submitted (19-Oct-1994) H. Lomholt, Dept. of Medical Microbiology,												
	The Bartholin Building, University of Aarhus, 8000 C Aarhus C,												
	DENMARK												
	1. location/Qualifiers												
	/organism="Neisseria meningitidis"												
	/strain="HF48"												
	/db_xref="taxon:487"												
	1.1489												
	/gene="iga"												
	<1. >1489												
	/gene="iga"												
	/codon_start=1												

QY 1398 ccaaggtcgtatcagcgtggtgagcaggtatcagtcatttggatcagcagcagataa 1457
|||||
Db 8234 CCAAGGCAATTAAGTGGCGAGGTACGGTTATCTGATCAACAAAGCCGATGCCGA 8175
QY 1458 aggcataaaacagccttctgtaaatcgtctgntcagcgccaggggtacggtgcaact 1517
|||||
Db 8174 CAAAAAAGTCACAGGCTTTCTCCCAAGTCGCAATGTGCAAGCGACGGGTACATTTGGTATT 8115
QY 1518 gaatgcgataatcagtcacacccgacaaactctatcttcgcttcgctgcgagcgttt 1577
|||||
Db 8114 AATTAATCTCAATACAGTATTAATCCGATTAACCTATTTGCGTTCCGTGGCGGTCGTTT 8055
QY 1578 ggaattaaacggagcattcgcttcgttcacacggtatcacaataacgagtgagggcgat 1637
|||||
Db 8054 GGATGCCAATGACATGACTTGTGAAACACATCCGCAACGTCGATGGAAGCGCGCG 7995
QY 1638 gattgncatcataatgacacacacatccacgcttacacatcaggaatgaagat 1697
|||||
Db 7994 CATTGTCAACCAACACAGGCCAGCGCTCCACATCCAGTAAAGGTTAATCTTTGAT 7935
QY 1698 tacacacacgagtggttaagaatalcaatagacttaattacagcaaga 1745
|||||
Db 7934 TACCGATCCCAAAACCATCTCTATTATTATTCATAAATATGATGA 7887
RESULT 33
AX044030 349980 bp DNA linear PAT 24-NOV-2000
LOCUS Sequence 109 from Patent W00066791.
DEFINITION AX044030
ACCESSION AX044030
VERSION AX044030.1 GI:11342914
KEYWORDS
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis.
bacteria: Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 349980)
AUTHORS Piazza,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C.,
Masimhan,V., Galeotti,C., Mora,M., Ratti,G., Scariselli,M.,
Scariato,V., Rappuoli,R., Frazer,C.M. and Grandi,G.
TITLE Neisseria genomic sequences and methods of their use
JOURNAL Patent: WO 0066791-A 109 09-NOV-2000.
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
Source Location/Qualifiers
I .349980
/organism="Neisseria meningitidis"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces.-seq 1: 1 to
349980 349980 bases-seq 108: 300001 to 649980 349980
bases-seq 109: 600001 to 949980 349980 bases-seq 110:
900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980
349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq
113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to
2272325 172325 bases"
BASE COUNT 82523 a 82940 c 96712 g 87805 t
ORIGIN
Query Match 4.4%; Score 191.4; DB 6; Length 349980;
Best Local Similarity 58.8%; Pred. No. 1.2e-32;
Matches 346; Conservative 0; Mismatches 239; Indels 3; Gaps 1;

Db 129928 AGGTGGGGGATTTGATGTTGCCAGCAAAAAGTCTTTGGCAAGTCAAAAATCCGAA 129869
QY 1338 aaagcagcgttccaaaatcggtcgaaggcagctgcaagttcaagcgaagggtgaaa 1397
|||||
Db 129868 TGGCGACAGATTGCGCAAAATCGCAAAAGGCACATTGAATAAACGCACAGCGTTAA 129809
QY 1398 ccaaggtcgtatcagcgtggtgagcaggtacagtcatttggatcagcagcagataa 1457
|||||
Db 129808 CCAAGGCAATTAAGTGGCGACGACGATGCGTTATCTGATCAATCAACAGCCATGCCGA 129749
QY 1458 aggcataaaacagccttctgtaaatcgtctgntcagcgccaggggtacggtgcaact 1517
|||||
Db 129748 CAAAAAAGTCACAGGCTTTCTCCCAAGTCGCAATGTGCAAGCGACGGTACATTTGGTATT 129689
QY 1518 gaatgcgataatcagtcacacccgacaaactctatcttcgcttcgctgcgagcgttt 1577
|||||
Db 129688 AATTAATCTCAATACAGTATTAATCCGATTAACCTATTTGCGTTCCGTGGCGGTCGTTT 129629
QY 1578 gattaaacggagcattcgcttcgttcacacggtatcacaataacgagtgagggcgat 1637
|||||
Db 129628 GGATGCCAATGACATGACTTGTGAAACACATCCGCAACGTCGATGGAAGCGCGCG 129569
QY 1638 gattgncatcataatgacacacacatccacgcttacacatcaggaatgaagat 1697
|||||
Db 129568 CATTGTCAACCAACACAGGCCAGCGCTCCACATCCAGTAAAGGTTAATCTTTGAT 129509
QY 1698 tacacacacgagtggttaagaatalcaatagacttaattacagcaaga 1745
|||||
Db 129508 TACCGATCCCAAAACCATCTCTATTATTATTCATAAATATGATGA 129461
RESULT 34
NM16139 1489 bp DNA linear BCT 14-MAR-1995
LOCUS NM16139
DEFINITION N.meningitidis iga gene (81139).
ACCESSION X82477.1
VERSION X82477.1 GI:732856
KEYWORDS iga gene; iga1 protease.
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis.
bacteria: Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 1489)
AUTHORS Lomholt,H., Poulsen,K. and Killian,M.
TITLE Comparative characterization of the iga gene encoding iga1 protease
JOURNAL in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
influenzae
Mol. Microbiol. 15 (3), 495-506 (1995)
MEDLINE 95302961
REFERENCE 2 (bases 1 to 1489)
AUTHORS Lomholt,H.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-1994) H. Lomholt, Dept. of Medical Microbiology,
The Bartholin Building, University of Aarhus, 8000 C Aarhus C,
DENMARK
FEATURES
Source Location/Qualifiers
I .1489
/organism="Neisseria meningitidis"
/strain="81139"
/db_xref="taxon:487"
/gene="iga"
1 .1489
<1 .>1489
/gene="iga"
/codon_start=1
/transl_table=1
/product="iga1 protease"
/protein_id="CAA57860.1"
/db_xref="GI:732857"
/db_xref="SPTREMBL:O51163"
/translation="ATDLSYKNGKGNHGDVSEDEYRSVQNDYEPKMKHNGHOGREDYDMA
VSVKHYAGDEISYTGHHNGHGDVSEDEYRSVQNDYEPKMKHNGHOGREDYDMA
RLNFEVTEVAPIAFLPSAGGVETTKDKNRSEFPAVVGAGTOPEYRSRNMTELSAYR

QY 1698 tacacaaccga 1708
| | | | |
Db 1455 TACCAATCCGA 1465

RESULT 36
NMIGAC16 1525 bp DNA linear BCT 14-MAR-1995
LOCUS N.meningitidis iga gene (NGC16).
DEFINITION X82485
ACCESSION X82485.1 GI:732869
VERSION iga gene; IgaI protease.
KEYWORDS Neisseria meningitidis.
SOURCE Neisseria meningitidis
ORGANISM Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.

REFERENCE 1 (bases 1 to 1525)
AUTHORS Lomholt, H., Poulsen, K. and Kilian, M.
TITLE Comparative characterization of the iga gene encoding IgaI protease
in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
influenzae
JOURNAL Mol. Microbiol. 15 (3), 495-506 (1995)
MEDLINE 95302961
REFERENCE 2 (bases 1 to 1525)
AUTHORS Lomholt, H.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-1994) H. Lomholt, Dept. of Medical Microbiology,
The Bartholin Building, University of Aarhus, 8000 C Aarhus C,
DENMARK

FEATURES
SOURCE Location/Qualifiers
1..1525
/organism="Neisseria meningitidis"
/strain="NGC16"
/db_xref="taxon:487"
1..1525
/gene="iga"
<1..>1525
/gene="iga"
/codon_start=1
/transl_table=11
/product="IgaI protease"
/protein_id="CA57868.1"
/db_xref="GI:732870"
/db_xref="SPTREMBL:O51167"
/translation="ATDLSYKNQGNIGNALSNVPMIDESVADYVKRIATVPQYA
VSVKAKAEVHTFYGOYNGHNDVADKENERVVEONNTEPHKMGASLQLEDITM
ARENKVTEVAAPLAPDAGGLDPTDKNRESSLRVGAGQOLYEEKAYHQEENKQ
YDLRDSOARVAYIAGTPYKQINDQTMTEGLIGFNHNTYSAEELKQALSDQALT
NGVLDGSGSLFAPDKQKNQWELGTYDWAAGYKSKSMOEMNITKKEFADKIKQRDN
ACTIKNGEHHNNTTGTNSHISTAVRAGNERDANNCONTEDEDNCTVLADQINQ
GAGGLFFKGDITYKGANNDITWLAGAGIDVADGKYYVQYKPNDRDLAKIKGTLEIN
GTGVNQGQLKANGDGTVLNOKADSNOYQASVQVIGVSGRTVLNNSNQINPDLVIF
GFGGRIIDANGNDLTFEHIRNVDEGARIVNHTQASITTLTGLSLINDPTTISLHYI
"

BASE COUNT 504 a 322 c 336 g 363 t
ORIGIN

Query Match 4.4%; Score 191.2; DB 1; Length 1525;
Best Local Similarity 60.3%; Pred. No. 6.5e-33;
Matches 332; Conservative 0; Mismatches 216; Indels 3; Gaps 1;

QY 1161 ttctatcgatagcgcaagcgaactcatctatcaacaacatcaaccaagcgcgagg 1220
| | | | |
Db 951 TGTACCTTTAAGACCAACCTTGTGTATTTGATCAAAACATCAACCAAGCGCGG 1010
| | | | |

QY 1221 cggcttgcatttgaagtgatttgaagtcgtcgctgaaacaacga--aacgtcgca 1277
| | | | |
Db 1011 CGGCTCTGTTTTCAAAGCGATTTACACAGTCAAAGTGCAAATTAATGACATCTGTCGT 1070
| | | | |

QY 1278 aggcgcggcgcttcatcaagtgaaagcagtcacgttacttgcgaagtaaacggcgctgc 1337
| | | | |

Db 1071 AGGTGCGGGGATTGATGTTGCCGACGCAAAAAAGCTGTTGGCAATCMAAAATCCGA 1130
QY 1338 aaagcagccgtgtccaaatattggcaaaagcagcgtgcacgttcaagccaaagggaa 1397
| | | | |
Db 1131 TGGCGACAGATTGGCAAAATCGGCAAAAGCACATTGAAATTAAGCGCACAGCGTTAA 1190
| | | | |

QY 1398 ccaaggtcgcattcagcgtgtgacagcagtcacgttcatcttggatcaagcagcagatata 1457
| | | | |
Db 1191 CCAAGGGCAATTAAGTACGCGGACGCTACGTTATTTCGAATCAAAAAACCGATTCAAA 1250
| | | | |

QY 1458 aggcacaaaacaagcccttagtgaatcggtcgttgcacgcgaggggtcagtgcaact 1517
| | | | |
Db 1251 CCAAAAAGCTCAGGCTTTCTCCCAAGCTCGCATTTGTCAGCGGACGCGCATTTGATTT 1310
| | | | |

QY 1518 gaatcgcaataatcggttcaacccgcgacaacactctattcggtcttcgcgcgagctt 1577
| | | | |
Db 1311 AATATGTTCAAAATCAGATTATATCCGATTAATCTTATTTGGTTTCGTTGGCGTGTCTT 1370
| | | | |

QY 1578 ggaattaaacgggacgttcgttcgttccacgcgtlatcaaaatacgaigaaggcgcat 1637
| | | | |
Db 1371 GGATGCCAATGGCAATGACATTGACTTTTGAAACATCCGCAAGCGTGAAGGCGCGC 1430
| | | | |

QY 1638 gattgncatcataatgccaacaacacatccacgcgttaccattacagggatgaagat 1697
| | | | |
Db 1431 CATTTGCAACACACACACAGCGCCACGCTCCACATCAAGCTAACGGGTAAATCTTGAT 1490
| | | | |

QY 1698 tacacaaccga 1708
| | | | |
Db 1491 TACCGATCCCA 1501

RESULT 37
AF012206 2992 bp DNA linear BCT 30-JAN-2001
LOCUS Neisseria meningitidis strain 24024 IgaI protease precursor (iga)
DEFINITION gene, partial cds.
ACCESSION AF012206
VERSION AF012206.1 GI:2529495
KEYWORDS
SOURCE
ORGANISM Neisseria meningitidis.
Neisseria meningitidis
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.

REFERENCE 1 (bases 1 to 2992)
AUTHORS Morelli, G., Malorny, B., Muller, K., Sella, A., Wang, J.-F., del
Valle, J. and Achtman, M.
TITLE Clonal descent and microevolution of Neisseria meningitidis during
30 years of epidemic spread
JOURNAL Mol. Microbiol. 25 (6), 1047-1064 (1997)
MEDLINE 98010345
REFERENCE 2 (bases 1 to 2992)
AUTHORS Morelli, G., Malorny, B., Mueller, K., Sella, A., Wang, J.-F., del
Valle, J. and Achtman, M.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-1997) Max-Planck Institut fuer Molekulare
Genetik, Innestrasse 73, Berlin 14195, Germany

FEATURES
SOURCE Location/Qualifiers
1..2992
/organism="Neisseria meningitidis"
/strain="24024"
/db_xref="taxon:487"
/note="serogroup A, subgroup VI; isolated in East-Germany,
1985"
1..>2992
/gene="iga"
/product="IgaI protease"
<1..>2992
/gene="iga"
/allele="iga6"
<1..>2992
/gene="iga"
/codon_start=3
/transl_table=11

mat_peptide
gene
CDS

/product="Iga1 protease precursor"
/protein_id="AC45789.2"
/db_xref="GI:1259796"
/translation="LAVALTPESEALVLRDVPDYQIFRDEAFENKGFVGAADPLSVN
KOGSNGLNSVPMIDSVADYNNRRITLIDPOYAVGVHVAGDETSYGGHNGHID
VSDENREYSVAONDIEPKNNHHNGREIDNMAINKFVTEVPIAFPTSAGGYE
TKDKRSEFVAVGADTOETNSRYNMLSLAVAYALAGTAYQDVNVSINQDEL
IGQDNKSHHSPERKLEVLSONALVAVLSDGSPLEADKOEKRWFLGAVADYVAG
YKNSQWENNIYKKEFEADEIKORDNAGTJIKGYEHWKTGTSTSTAVRAGNR
DANQONVTEFEDNTVLDPONINOGAGJLEFGDYVYKANNITMIGADIVADCK
VWVQNPNGDRLAKIGKGTLEINGVNOGGLKVGDTYILNOKADSNOKYVAFQV
GYSGHTLVINSNOINPDNIYFEGFGRGLDANGDYFEHILANDDEARITNHS
HASTITLCKSLITAPONLSYEIRNDYDDDDYGYSTYKPIPOCKDLYKRYRIA
LKSQSVNAPMEPNVTEENDVFEKGTDEAKNNMNNRISGFSEFGEENK
KHGALNLNENKSAONRFLITGGALNGLISYQGVLLISGPRTPHARDFVKSAR
KDAHFENNEVEEDWIRNTEKAAITVQVSAFSSGRVSDITANITADNAKVL
GYKNGDEVCRSDTYGVCNTDNLSDKMLNSGALQINGNVMSALVILKALM
GOIOGNSRVSLNOSHMLTGDQVHNLSDSHILNLSADQASANKYHYTLKIH
LSGNGEHLTHLAKNLGDKVLYKESAGHYLODKGEPNDOGLNFDASVODR
SRLSVSLANNHDLGALRYTKITENGITRLINRYANRRKPAFPAITNTAQAOKA
TQTDGQRIKPNQIVYAPPSPOAN"

BASE COUNT 983 a 684 c 639 g 686 t
ORIGIN

Query Match 4.4%; Score 191.2; DB 1; Length 2992;
Best Local Similarity 60.6%; Pred. No. 7,1e-33;
Matches 329; Conservative 0; Mismatches 211; Indels 3; Gaps 1;

1161 ttatcatgattacgcaacgcaaacatcatctatacaacaataccaagcgcg 1220
1025 TGTACCTTGAAGACACGACCTTGTGATTCGATCAAAACCAACGCGCGG 1084
1221 cggttgatttgaagagtgattacagtcctcgccgtaaacaa---cgaaacgtgaca 1277
1085 CCGTCTCTTTTTCAGAGCGCATACAGTCAAGTGCATTAATGCGATCCTTGCT 1144
1278 aggcgcggcgcttcatacatcagtgaaagatccgcttacttgaaaglaaacggcgatgac 1337
1145 AGGTGCGGGGATTGATGTTCCCGACGCAAAAAGTCGTTGGCAAGTCAAAAATCCGAA 1204
1338 aaacgacgccttcacaatatcgcaagcgacgtcgacgttcacgaagcggaagaa 1397
1205 TGGCGACAGATTCGCAAAAATCGGCANAGGCACACTTGAATAAAGCGCACAGCGCTTAA 1264
1398 ccaagcgcgacgacgctgagcgagtgacagtcatttgatcagcagcagcagcagcagcagc 1457
1265 CCAAGGCGCAATTAAGTCCGCGACGCTGATTCGTAATCAAAAAGCGCATTCAAA 1324
1458 aggcacaaacacagccttagtaaatcgcttgnatcagcgacgagcgagtgatgacgaact 1517
1325 CCAAAAAGTCCAGGCTTCTCCCAAGTCGGCATTTGCAAGCGGCGGTACATTGATTT 1384
1518 gaatgcgataatcagtcacacccgcaaacatctatcttcggtcttcgcgcgacgcttt 1577
1385 AATGATGTTCAATACAGATTAATCCCGATTAATTTCCGTTCCGTTCCGTTCCGTTCCGTT 1444
1578 gggatttaaacggcattcgcttcgctcaccgctatcaaaataccgataagggcgat 1637
1445 GGAATGCCAATGCAATGACTTTGTAACACATCCGACATCGGATGAAGCGCGCGC 1504
1638 gattgcnatcataatcccaacaacatccacgcttaccatcaaggaagaaatgaat 1697
1505 CATTTGCAACCAACCAACCAAGCGCCTCCACATACATTAACGCGCAATATTAAT 1564
1698 tac 1700
1565 TAC 1567

RESULT 38
NGIGAPRO 4899 bp DNA linear BCT 21-MAR-1995
LOCUS

DEFINITION Neisseria gonorrhoeae gene for Iga protease.
ACCESSION X04835
VERSION X04835.1 GI:44868
KEYWORDS Iga protease; inverted repeat; protease; secreted protein; signal peptide.
SOURCE Neisseria
ORGANISM Neisseria gonorrhoeae.
Neisseria gonorrhoeae
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 4899)
AUTHORS Pohlner J., Haler R., Beyreuther K. and Meyer T.F.
TITLE Gene structure and extracellular secretion of Neisseria gonorrhoeae Iga protease
JOURNAL Nature 325 (6103), 458-462 (1987)
MEDLINE 87115823
FEATURES
source location/Qualifiers
1..4899
/organism="Neisseria gonorrhoeae"
/strain="M11"
/db_xref="taxon:485"
/clone="p1p100"
34..39
/note="put. -43 region"
43..48
/note="put. -35 region"
63..68
/note="put. -10 region"
87..97
/note="put. rRNA binding region"
104..4702
/note="Iga protease precursor; 169K (AA -27 to 1505)"
/codon_start=1
/transl_table=11
/protein_id="CA48538.1"
/db_xref="GI:44869"
/db_xref="SWISS-Prot:P09790"
/translation="MKARRKINLSITSLATLPTSEALVLRDVPDYQIFRDEAF
NKRFVGAADPLSVNKNRGONIGALSVMPIDESADVKRIATVDPQAVSKRA
KAEVHTFYQYNGHNDVADKEMERYEEDNNNEPHAWKASNLGRLEDYNNMAFNF
VTEVAPVIACTDAGGDLTYKDKNFSFVRIAGROLYEKGVYHOGSNKGVLDRL
SOAVRYAIACTPKDINIADOTMTEGLIGGNHKOYSAEELKQSLDQALITNGVLG
DCSPLFAFPKONQWFLGTVDYMAQYKNSQWENNIYKKEFADKIKODNAGTVG
NGEHHKTTGTNSHISTAVRIANMEGADNNGCNVPEPDGTTLVNONTNNOGAGLEF
KGYTYGANNITWLAGIDVADGKRYVQYKRPNDRLAKIGKGTLEINGTVNG
QLVGDVTLINORADKRVQVAFSGVITSGRTLVLSNQINPNLYFEGFGR
DANGNDLTFEHLINVEDGARIVNHNIDHASTITLTKSLITNPNSLVHSIONDYED
DYSYYRPRRPIQGGKDLVRYKRYLKSGLRANMPENGVAENDWIFMGYTOE
ARKNAMHKNRIRIGDGFDEENGSGHNGALNENKSAOKRFLITGGALNGLKI
SVQGNVLLSGRPHARDFVKNSSAKDAHFESKNNEVPEDDMIINTERFAAETAVNO
SASFSGRNVSIDITANTTATDNMKVNLGYRNGDEVCRSDTYGVCNTGNLSKALN
SEFATRLNGVNLNQNALVLCRAALMGKTIQCGNSVSLNOSHMLTGDQVHNL
LADSHILNNSADQASANKYHTITLNSGNGHHTLTLAKNLGDLVLYKESAGHY
QLHVQNTGEPNDEGLDLPDASSVQDSRLVSLANVYDLGALRYIKENGTTRYL
NPYAGNRPVYKPAFPAANASQAKATQDGAQIAKPOINIVAPSPPOANOAEALR
QOAAKEQVNRQOAAEAERVARODEAKRAAEIARQOEBAKAEIARQOAAEAERK
AREIARQAEASIOAAKAPKRRRRRIILRPAPVPSLDDYDAKDNSESSIGLARV
IPWGRELIINDYEELPLELEDEAREERORLOPHSKSRNRRAISPPSSDEASAV
STSDKRPHODNTEHEKYETAGLQPRAAOPTQOAAQADAVSTNNSLSAMASTOI
LSDMGATLTHIAOKSRADAEKNSVMSNVTGIDIASQIRRSKRTOIQTIDS
LBNMGIGVLTYSDSQHTPDQAGKNTFQVQNLXKYYLNDAYVAGDIQAGSLRS
LQTOOKANFNRTSIOTGLTGLNPLKINQFETVPSAGIRYRLSADYKLDQSVKSS
MAVKTLAGIDFAVRFKVGNLVYKPLISAAVFAVNGGCVNGGKSPAYRADNQQOYS
AGVALLRYNTLVNNGISITKQOLEKQSQIKQIRF"

sig_peptide 104..184
mat_peptide 185..4699
misc_feature 3061..3062
/note="mature Iga protease (AA 1-1505)"
misc_feature 3157..3158
/note="autoproteolytic site for Iga protease"
misc_feature 3466..3467
/note="autoproteolytic site for Iga protease"
repeat_unit 4725..4739
/note="autoproteolytic site for Iga protease"

1221 cggcttgatattgaagtgatttaccgctcgcctgaaacaacga---aacgtggca 1277
1267 CGGCGCTTTTTCOAAGGCGATTACACAGTCOAAGGGCGCAATATGACATCACTTGGTT 1326
1278 aggcgcggcgcttcaatcaatgaagaacgtagccgttaacttggaaagaagcgcgtgac 1337
1327 AGGTGGCGGGGATTGATTGTTCCGCGCAAAAAGTGTGTTGGCAAGTCAAAATCCGAA 1386
1338 aaacgcgcgccttccaaatcgcaaaagcagctgacgttcaagcgaaggga 1397
1387 TGGCGACAGATTGGCAAAATCGCANAAGCGCTTTGGAAATTAACGGCACAGCGCTTAA 1446
1398 ccaagcctcagcagcgtggcgacggtacagtcatttggatcagcagcagaagataa 1457
1447 CCAAGGCGCAATTAAGGTGGCGGCGGTACGGTTATTCGTGATCAAAAAGCGGATCCGA 1506
1458 aggcgaacaaacagccttgaatcgagctgntcagcgagcgaggtgacgtgcaact 1517
1507 CAAAAAAGTTCAAGCTTTCCCAAGTCGGCATGTGTAGCGGCGCGGTACATTGTATT 1566
1518 gaatgcgataatcagtcaccccgcaaaactatatttggcttgcggcgagcgttt 1577
1567 AATTAAGTCAATCAGATTATATCCGATTAACCTATATTTCGTTTCGTCGGCGCTT 1626
1578 ggaattaaacggcagcagcttgcctccacgcatcaaatcagcagtgaaagcgcat 1637
1627 GGTATGCCAATGGCAATACCTTTTGAACACATCCGCANACGTGATGAAGGCGCG 1686
1638 gattgncatcaataatgccaacaacatccacggttaaccattcaggaaggaatgaagat 1697
1687 CATTGTCAACCAACACACAGACCGCTCCACATACACTTAACGGGTAAATCTTTAAT 1746
1698 tacacaacga 1708
1747 TACCAATCCGA 1757

RESULT 42
NGCOMLGN 8393 bp DNA linear BCT 18-NOV-1996
LOCUS N.gonorrhoeae gene for competence lipoprotein ComL.
DEFINITION 249895
ACCESSION 249895.1 GI:1107832
VERSION 1
KEYWORDS ComL; competence lipoprotein.
SOURCE Neisseria gonorrhoeae.
ORGANISM Neisseria gonorrhoeae;
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
1 (bases 1 to 8393)
Fussenegger, M., Facius, D., Meier, J. and Meyer, T.F.
A novel peptidoglycan-linked lipoprotein (ComL) that functions in
natural transformation competence of Neisseria gonorrhoeae
Mol. Microbiol. 19 (5), 1095-1105 (1996)

JOURNAL MEDLINE 96249702
REFERENCE 2 (bases 1 to 8393)
AUTHORS Fussenegger, M.
TITLE Direct Submission
JOURNAL Submitted (20-JUN-1995) Martin Fussenegger Dr., Infektionsbiologie,
Max-Planck-Institut, fuer Biologie, Speemannstr. 34, Tuebingen, BW,
D-72076, Germany
location/Qualifiers
1. .8393
/organism="Neisseria gonorrhoeae"
/strain="MS11A"
/db_xref="taxon:485"
120. .923
/codon_start=-1
/transl_table=-1
/product="ComL, competence lipoprotein"
/protein_id="CAA90076.1"
/db_xref="GI:1107833"
/db_xref="SPTREMBL:O50985"
/translation="MKRLITVSLGIALSACATGTADKDAQTODWSVEKLYAEADQ

ELNSNMYTRAVKLYELLERSFPTSRHAROSQOLDPAVAYYKXDEKALAIERFRH
POHPMDYALYRGVLFPNEDOSFLNKLASODMSDDPKANREAYQAFALVORFNS
KYAADTARWVKLVDAIGNEMSRGAYTYMKGAYTIAANAKRLTIGSOYTRVEESL
ATLEAYKKRLDPQLAADTRVLETFNFPSPFLTHMQPDMMPMRWH"

terminator 934..967
BASE COUNT 2550 a 2300 c 1886 g 1657 t
ORIGIN

Query Match 4%; Score 191.2; DB 1; Length 8393;
Best Local Similarity 60.3%; Pred. No. 8.le-33;
Matches 332; Conservative 0; Mismatches 216; Indels 3; Gaps 1;

1161 ttattcgattcagcaagcgaacactcattcacaacaacacccaagcgcg 1220
4701 TGTACTCTTTGAAGACACACGCACTTGGTATTGAATCAAAACATCAACGACGCGCG 4760
1221 cggcttgatattgaagtgatttaccgctcgcctgaaacaacga---aacgtggca 1277
4761 CGGCTGTTTTCOAAGGCGATTACACAGTCOAAGGCGCAAAATATGACATCACTTGGTT 4820
1278 aggcgcggcgcttcaatcaatgaagaacgtagccgttaacttggaaagaagcgcgtgac 1337
4821 AGGTGGCGGGGATTGATTGTTCCGCGCAAAAAGTGTGTTGGCAAGTCAAAATCCGAA 4880
1338 aaacgcgcgccttccaaatcgcaaaagcagctgacgttcaagcgaaggga 1397
4881 TGGCGACAGATTGGCAAAATCGCANAAGCGCTTTGGAAATTAACGGCACAGCGCTTAA 4940
1398 ccaagcctcagcagcgtggcgacggtacagtcatttggatcagcagcagaagataa 1457
4941 CCAAGGCGCAATTAAGGTGGCGGCGGTACGGTTATTCGTGATCAAAAAGCGGATCCGA 5000
1458 aggcgaacaaacagccttgaatcgagctgntcagcgagcgaggtgacgtgcaact 1517
5001 CAAAAAAGTTCAAGCTTTCCCAAGTCGGCATGTGTAGCGGCGCGGTACATTGTATT 5060
1518 gaaatgcgataatcagtcaccccgcaaaactatatttggcttgcggcgagcgttt 1577
5061 AATTAAGTCAATCAGATTATATCCGATTAACCTATATTTCGTTTCGTCGGCGCTT 5120
1578 ggaattaaacggcagcagcttgcctccacgcatcaaatcagcagtgaaagcgcat 1637
5121 GGTATGCCAATGGCAATACCTTTTGAACACATCCGCANACGTGATGAAGGCGCG 5180
1638 gattgncatcaataatgccaacaacatccacggttaaccattcaggaaggaatgaagat 1697
5181 CATTGTCAACCAACACAGACCGCTCCACATACACTTAACGGGTAAATCTTTAAT 5240
1698 tacacaacga 1708
5241 TACCAATCCGA 5251

RESULT 43
NMIGA341 1489 bp DNA linear BCT 12-JUL-1999
LOCUS N.meningitidis iga gene (IHNS341).
DEFINITION X82478
ACCESSION X82478.1 GI:732862
VERSION 1
KEYWORDS iga gene; iga1 protease.
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis;
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
1 (bases 1 to 1489)
Lomholt, H., Poulsen, K. and Killian, M.
Comparative characterization of the iga gene encoding iga1 protease
in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
influenzae
Mol. Microbiol. 15 (3), 495-506 (1995)

JOURNAL MEDLINE 95302961
REFERENCE 2 (bases 1 to 1489)

Db 1215 CAAAAAGTCAGGCTTTCTCCCAAGTCAGGATGTCAGCGGAGCCGTACATTGTTAT 1274
Qy 1518 gaatgccgaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat 1577
Db 1275 AATTAATTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1334
Qy 1578 ggaattcaacggcgatcgttcgttcacacggtatcaatcaatcaatcaatcaatcaat 1637
Db 1335 GGATGCCAATGGCAATGACTTGTGACACATCCGACAGCGTGAATGAAGCGCGCG 1394
Qy 1638 gattgcnatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat 1697
Db 1395 CATTTGCAAC 1454
Qy 1698 taccaaccga 1708
Db 1455 TACCGATCCCA 1465

RESULT 45
NMIGAC65
LOCUS NMIGAC65 1525 bp DNA linear BCT 14-MAR-1995
DEFINITION N.meningitidis i9a gene (NGC65).
X82484
VERSION X82484.1 GI:732871
KEYWORDS i9a gene; Iga1 protease.
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.

REFERENCE 1 (bases 1 to 1525)
AUTHORS Lomholt, H., Poulsen, K. and Kilian, M.
TITLE Comparative characterization of the i9a gene encoding Iga1 protease
in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
influenzae
JOURNAL Mol. Microbiol. 15 (3), 495-506 (1995)
MEDLINE 95302961
REFERENCE 2 (bases 1 to 1525)
AUTHORS Lomholt, H.
TITLE Direct Submision
JOURNAL Submitted (19-OCT-1994) H. Lomholt, Dept. of Medical Microbiology,
The Bartholin Building, University of Aarhus, 8000 C Aarhus C,
DENMARK

FEATURES
source Location/Qualifiers
1..1525
/organism="Neisseria meningitidis"
/strain="NGC65"
/db_xref="taxon:487"
1..1525
/gene="iga"
<1..>1525
/gene="iga"
/codon_start=1
/transl_table=11
/product="Iga1 protease"
/protein_id="CA57867.1"
/db_xref="GI:732872"
/translation="ATDLSYKKNKQONIGNALSNPMDISVADVNRKRIATVVDPOA
VSUHKAKAEVTFITGYNGHNDVADKENEYRVQNNTEPHKAMSASNLGRLEDYNM
ARFKFVTEVAPLAPFDAGGLDLYKDKNRFSPFVAVGAGRLVVEKGAVHOGENEG
YDLRDLQAVRYALATGPKDINIDOTMTEGLIGGNHNTYSAELKQALSODALT
NRYVLDGSGSLPFAADKQKQNWVFLGTVDYAGYGRKSWQEWNIYKKEPDKIKORDN
AGTILKNGEHHNMTFFGNSHIGSTAVRLAGNEKDNANGONVTFEDNGTLVIDONIN
GAGGLEFGDVTGKANDITMLGAGIDVADGKVVWQVKNPNDRLAKIGKCTLEIN
GTGVNCGGLKYGDGVILLNOKADSNQKYQAFSQVIGVSGRTLVLSNQINDNLIF
GFRGGRIDANGNDLTFEHRINRYDEGARIVNHTGASTITLTGKSLITAPQNLVYEI

BASE COUNT 510 a 320 c 330 g 365 t
ORIGIN

Query Match

4.4%; Score 189.6; DB 1; Length 1525;

Best Local Similarity 60.4%; Pred. No. 1,5e-32;
Matches 328; Conservative 0; Mismatches 212; Indels 3; Gaps 1;

Qy 1161 ttltatcgattacggaacggaacacatctatcaataacacatcaacgaagcgcg 1220
Db 951 TGTTCACCTTTGAAGACACACGCGCTTTGGTATTGGATTCAAACATCAACACGCGCGG 1010
Qy 1221 cgggttgatcttgaagtgatcttcaaggtctcgctgaacaacga---aacgtgca 1277
Db 1011 CGGTCTGTTTTTCMAAGGGGATTCACAGTCMAAAGGTCMAAATATGACATCTTGCT 1070
Qy 1278 aggcgcgcggttcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat 1337
Db 1071 AGTGGGCGGATTTGATGTTGCCGACGCGCAAAAAGTCGTTGCAAGTCAAAAATCCGAA 1130
Qy 1338 aacgacgcgctgtccaaaatcgcaaaagcagcgtgcacgttcacgaacgaaggggaaa 1397
Db 1131 TGGTACACAGATTGGCAAAAATCGGCAAGGACACATTAGAAATTAACGGCACACGCGTTAA 1190
Qy 1398 ccaaggtcgcatacagcgtgagcagcgtacagtcatttgatcagcagcagcagataa 1457
Db 1191 CCAAGGGCAATTAAAGTGGCGGACGACAGTTATTGTAATCAAAAAGCCGATTCMAA 1250
Qy 1458 agycaaaaacgaagcctttagtgaatcagcgttgcagcgcgagcgtacggtgcaact 1517
Db 1251 CCAAAAAGTCACAGGCTTTCTCCCAAGTCGCGCATTTGTCAGCGGACGCCGTCATTGCTATT 1310
Qy 1518 gaatgcgaataatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat 1577
Db 1311 AATATGTTCAAAATACGATTAATTCGATTAATTCGATTAATTCGATTAATTCGATTA 1370
Qy 1578 ggaattcaacggcgatcgttcgttcacacggtatcaatcaatcaatcaatcaatcaat 1637
Db 1371 GGATGCCAATGGCAATGACTTGTGACACATCCGACAGCGTGAATGAAGCGCGCG 1430
Qy 1638 gattgcnatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaatcaat 1697
Db 1431 CATTTGCAAC 1490
Qy 1698 tac 1700
Db 1491 TAC 1493

Search completed: June 30, 2002, 12:42:41
Job time: 69979 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: July 1, 2002, 02:30:47 ; Search time 415.2 seconds
(without alignments)

528.893 Million cell updates/sec

Title: US-09-303-518D-571

Perfect score: 894
1 atggttcgttaccatcag.....accgcataaacgcgtaa 894

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 38353 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 100 summaries

Database : Issued_Patents_NA.*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCMus_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38.6	4.3	4403765	4 US-09-103-840A-2	Sequence 2, Appl
2	37	4.1	4411529	4 US-09-103-840A-1	Sequence 1, Appl
3	35.8	4.0	7218	1 US-08-232-463-14	Sequence 14, Appl
4	35.2	3.9	4411529	4 US-09-103-840A-1	Sequence 1, Appl
5	34.2	3.8	285	4 US-09-007-005-17	Sequence 17, Appl
6	34.2	3.8	289	4 US-09-244-796-17	Sequence 17, Appl
7	34.2	3.8	2277	1 US-08-676-967-5	Sequence 5, Appl
8	34.2	3.8	2277	1 US-08-676-974-5	Sequence 5, Appl
9	34.2	3.8	2277	1 US-09-098-487-5	Sequence 5, Appl
10	34	3.8	1198	3 US-09-248-335-27	Sequence 27, Appl
11	33.8	3.8	2085	1 US-08-553-999B-1	Sequence 2, Appl
12	33.8	3.8	3358	3 US-09-248-571-2	Sequence 2, Appl
13	33.6	3.8	2191	3 US-09-035-555B-12	Sequence 12, Appl
14	33.6	3.8	35081	2 US-08-752-760A-1	Sequence 1, Appl
15	33.4	3.7	2070	1 US-07-766-351-6	Sequence 6, Appl
16	33.4	3.7	2070	1 US-08-059-032-6	Sequence 6, Appl
17	33.4	3.7	2070	5 PCT-US91-07290-6	Sequence 6, Appl
18	33.2	3.7	1155	4 US-09-422-487-4	Sequence 4, Appl
19	33.2	3.7	1158	3 US-08-844-188-45	Sequence 45, Appl
20	33.2	3.7	20235	1 US-07-642-734C-3	Sequence 3, Appl
21	33.2	3.7	20235	3 US-08-439-009A-3	Sequence 3, Appl
22	33	3.7	1119	4 US-08-987-943-2	Sequence 2, Appl
23	33	3.7	1134	3 US-09-248-335-29	Sequence 29, Appl
24	33	3.7	1755	4 US-08-987-943-1	Sequence 1, Appl
25	33	3.7	7898	4 US-08-984-709A-49	Sequence 49, Appl
26	32.8	3.7	1473	2 US-08-541-033A-25	Sequence 25, Appl
27	32.8	3.7	1473	2 US-08-828-451-25	Sequence 25, Appl

Db 606 tggcgcagctctccaccctgccaaatccaccataatcgtcaactctccgacaggcgca 665
Qy 464 tcaagagcatagacaataatcatgagcgagtgcgcgcaagaagcaaaaccgagc 523
Db 666 ggaagccttcaccgcaggaagacgctgagcgaagtgtacgaagaatctccaccgagc 725
Qy 524 ccaccgcatatacaagggttcaaacatcatcaaggccctgcgcgcggcgaggaaccca 583
Db 726 acccttcggaagcagtgatcatgaatgacagaaggagcaccgccgcgctgcatgtatgcg 785
Qy 584 tcatctgcgc 593
Db 786 tctgtcttc 795

RESULT 11
US-08-553-999B-1
; Sequence 1, Application US/08553999B
; Patent No. 5773599
; GENERAL INFORMATION:
; APPLICANT: Govindaswamy Chinadural
; TITLE OF INVENTION: Isolation and characterization of cDNA for E1A C-terminal bind
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: G. Chinadural
; STREET: 3681 Park Avenue
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 713 Kb storage
; COMPUTER: IBM Compatible (Gateway 2000)
; OPERATING SYSTEM: Windows 3.1
; SOFTWARE: Word for Windows (6.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/553,999B
; FILING DATE: 11/06/95
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5773599e
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2085 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: human
; INDIVIDUAL ISOLATE: PACT30
; CELL TYPE: B-cell
; IMMEDIATE SOURCE:
; LIBRARY: cDNA
; CLONE: PACT30
; PUBLICATION INFORMATION:
; AUTHORS: Ute Schaeper
; AUTHORS: Janice M. Boyd
; AUTHORS: Sulekha Verma
; AUTHORS: Erik Uhlmann
; AUTHORS: T. Subramanian
; AUTHORS: G. Chinadural
; TITLE: Molecular cloning and characterization of a cellular phosphoprotein that
; JOURNAL: Proceedings of the National Academy of sciences United States of Americ
; VOLUME: 92
; ISSUE: 23
; PAGES: 10667-10671
; RELEVANT RESIDUES IN SEQ ID NO: 1:

US-08-553-999B-1
Query Match 3.8%; Score 33.8; DB 1; Length 2085;
Best Local Similarity 47.8%; Pred. No. 2.2;
Matches 98; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
Qy 435 cctgaccgcctgtataaagccgcgaataatcaagatagaaatcatgtcagcggc 494
Db 1135 CATGACCCCGCCGCTGTGTCACCTGACCTCAAGCTCAATGAGGCTGCTATAGTACCTCCGG 1194
Qy 495 caaggtgcgcgcgaagaagcaaaacgcgcaccgcgcacatacaagggtcaaacaaatcat 554
Db 1195 CGTGGTGGCGGTGGCCGCCCATGCGATCCAGCTGCTGTGAAGTATCTCCCGAGCC 1254
Qy 555 caagccctgc 614
Db 1255 CATGTCCTGTCTCCACAGGCGCTGCCCTGTGGCCACCGCCCGCCCGCTTCTCTCG 1314
Qy 615 ggaagcgcgcgctgtgagcgat 639
Db 1315 CCAACCGTCAAGCCCGAGCGGAT 1339

RESULT 12
US-09-248-571-2
; Sequence 2, Application US/09248571
; Patent No. 6136539
; GENERAL INFORMATION:
; APPLICANT: BASBAUM, CAROL
; APPLICANT: GALLUP, MARIANNE
; APPLICANT: DAIZONG, LI
; APPLICANT: GEREMICHAEL, ASSEFA
; APPLICANT: GENSCH, ERIN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INHIBITION OF MUC-5 MUCIN
; FILE REFERENCE: USFT12/02
; CURRENT APPLICATION NUMBER: US/09/248,571
; CURRENT FILING DATE: 1999-02-11
; EARLIER APPLICATION NUMBER: 60/074,398
; EARLIER FILING DATE: 1998-02-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 3358
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-248-571-2

Query Match 3.8%; Score 33.8; DB 3; Length 3358;
Best Local Similarity 58.4%; Pred. No. 2.7;
Matches 59; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 345 ggaagaaggcggaaggctgtctgtatcatcagcgacatcgagctacgtattggcgg 404
Db 3002 ggaagagagccaggaagggtccatcacaccatccgcgcagatggcgtatccctgtgtga 3061
Qy 405 acgtatcatcagccagcagcttcgtccacttgacgcca 445
Db 3062 caccgacatggcctgtgtctgtcgtggaacaagaagacca 3102
RESULT 13
US-09-039-555B-12/c
; Sequence 12, Application US/09039555B
; Patent No. 6033856
; GENERAL INFORMATION:
; APPLICANT: Koerner, Kathrin
; APPLICANT: Mueller, Rolf
; APPLICANT: Sadlasek, Hans-Harald
; TITLE OF INVENTION: PROMOTER OF THE CDC25B GENE, ITS
; TITLE OF INVENTION: PREPARATION AND USE

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/766,351
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Lisabeth Felix
REGISTRATION NUMBER: 31547
REFERENCE/DOCKET NUMBER: 17796-002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 877-0900
TELEFAX: (415) 877-8370
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2070 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-07-766-351-6

Query Match
Best Local Similarity 47.4%; Score 33.4; DB 1; Length 2070;
Matches 100; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
QY 381 catcgcaactacgatttgaggcgagcgtacacagcagcagcttcggtccacctgac 440
DB 732 CAACCTCCCGGTGCAAGAGAGAGAACTGCGTATCTCTCAAGAGAGTGTGACGCTGGGGG 791
QY 441 cgcacgtacaaagcgccggaataatcaagagatagacaaatcatgcaaggcgagcaggt 500
DB 792 CCAGAAGTCCCGCTGCTGGGTTCCACACGACGCCGACTATGTCTGGAGATGAACAT 851
QY 501 ggcggcgcaaaagcgaacccgcccagcgccacacacacacacacacacacacacacacac 560
DB 852 GGCACAGACACCAACGACGCGTGGCCACCTTCTAGATGAGAGCTGGCGCAGAGCTGAAGCC 911
QY 561 cctgcgcggcgagcgagcaaccatcatcctg 591
DB 912 CTTGGGGGAGCAGAGCGCTGCGGTGATTCTG 942

RESULT 16
US-08-059-032-6
Sequence 6, Application US/08059032
Patent No. 5424205
GENERAL INFORMATION:
APPLICANT: Sinha, Sukanto
APPLICANT: Seubert, Peter A.
APPLICANT: Dovey, Harry F.
APPLICANT: McConlogue, Lisa C.
APPLICANT: Little, Sheila P.
APPLICANT: Johnstone, Edward M.
TITLE OF INVENTION: Amyloidin Protease and Uses Thereof
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend Kourlie and Crew
STREET: One Market Plaza, Steuart Street Tower,
STREET: Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/059,032

FILING DATE: 19930507
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William E.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 15270-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2070 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-059-032-6

Query Match
Best Local Similarity 47.4%; Score 33.4; DB 1; Length 2070;
Matches 100; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
QY 381 catcgcaactacgatttgaggcgagcgtacacagcagcagcttcggtccacctgac 440
DB 732 CAACCTCCCGGTGCAAGAGAGAGAACTGCGTATCTCTCAAGAGAGTGTGACGCTGGGGG 791
QY 441 cgcacgtacaaagcgccggaataatcaagagatagacaaatcatgcaaggcgagcaggt 500
DB 792 CCAGAAGTCCCGCTGCTGGGTTCCACACGACGCCGACTATGTCTGGAGATGAACAT 851
QY 501 ggcggcgcaaaagcgaacccgcccagcgccacacacacacacacacacacacacacacac 560
DB 852 GGCACAGACACCAACGACGCGTGGCCACCTTCTAGATGAGTGGCGCAGAGCTGAAGCC 911
QY 561 cctgcgcggcgagcgagcaaccatcatcctg 591
DB 912 CTTGGGGGAGCAGAGCGCTGCGGTGATTCTG 942

RESULT 17
PCT-US91-07290-6
Sequence 6, Application PC/TUS9107290
GENERAL INFORMATION:
APPLICANT: Sinha, Sukanto
APPLICANT: Seubert, Peter A.
APPLICANT: Dovey, Harry F.
APPLICANT: McConlogue, Lisa C.
APPLICANT: Little, Sheila P.
APPLICANT: Johnstone, Edward M.
TITLE OF INVENTION: Amyloidin Protease and Uses Thereof
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Athena Neurosciences, Inc.
STREET: 800F Gateway Blvd.
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/07290
FILING DATE: 19911004
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Lisabeth Felix
REGISTRATION NUMBER: 31547

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/541,033A
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI155
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1473 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 4..1431
US-08-541-033A-25

Query Match 3.7%; Score 32.8; DB 2; Length 1473;
Best Local Similarity 45.4%; Pred. No. 3.7; Indels 0; Gaps 0;
Matches 118; Conservative 0; Mismatches 142;

QY 285 ggaagacatgaaacaatgttcaaaagcgtgtacacgctggaaacgctgacagcgttt 344
DB 972 gTACAAGAGCAGACCCCGCTGATGTGGCGAGACCCCGAAGCCTTGGAGCTGACTG 1031
QY 345 ggaacagggggaagggctgtgttcatcaacgacgacatcgacgacatttgagcgg 404
DB 1032 CCAAGTGAATCGCTTCCCTGCGCCGACCAAGAGATCGATGAGCAGCGCGGA 1091
QY 405 acgcatatcagcagcagcgttcgcttcacacgtgacccgcatgtacagcgcgcaaat 464
DB 1092 GCTGCTGATCAAGCAGCGCTGCGCAGTACGTGTGGAGGGGCCAACAATGCCCTCCACCA 1151
QY 465 caaagcagatagacaataatcatgcaagcggcgaggggtgcggcgaaggaacacgcgc 524
DB 1152 CGAGGCCATCCACAAGTACAACAGGCCGCGATCATCTACTGCCCGCAAGCGGCCAA 1211
QY 525 caccgcatataaagggtca 544
DB 1212 CGCCGGCGCGGTGGCGGTCA 1231

RESULT 27
US-08-828-451-25
Sequence 25, Application US/08828451
Patent No. 5985634
GENERAL INFORMATION:
APPLICANT: Schmidt, Robert R.
APPLICANT: Miller, Philip
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES
RELATING TO THE - AND -SUBUNITS OF GLUTAMATE
TITLE OF INVENTION: DEHYDROGENASES AND METHODS OF USE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSER: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/828,451
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/541,033
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI155
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1473 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 4..1431
US-08-828-451-25

Query Match 3.7%; Score 32.8; DB 2; Length 1473;
Best Local Similarity 45.4%; Pred. No. 3.7; Indels 0; Gaps 0;
Matches 118; Conservative 0; Mismatches 142;

QY 285 ggaagacatgaaacaatgttcaaaagcgtgtacacgctggaaacgctgacagcgttt 344
DB 972 gTACAAGAGCAGACCCCGCTGATGTGGCGAGACCCCGAAGCCTTGGAGCTGACTG 1031
QY 345 ggaacagggggaagggctgtgttcatcaacgacgacatcgacgacatttgagcgg 404
DB 1032 CCAAGTGAATCGCTTCCCTGCGCCGACCAAGAGATCGATGAGCAGCGCGGA 1091
QY 405 acgcatatcagcagcagcgttcgcttcacacgtgacccgcatgtacagcgcgcaaat 464
DB 1092 GCTGCTGATCAAGCAGCGCTGCGCAGTACGTGTGGAGGGGCCAACAATGCCCTCCACCA 1151
QY 465 caaagcagatagacaataatcatgcaagcggcgaggggtgcggcgaaggaacacgcgc 524
DB 1152 CGAGGCCATCCACAAGTACAACAGGCCGCGATCATCTACTGCCCGCAAGCGGCCAA 1211
QY 525 caccgcatataaagggtca 544
DB 1212 CGCCGGCGCGGTGGCGGTCA 1231

RESULT 28
US-08-541-033A-23
Sequence 23, Application US/08541033A
Patent No. 5879941
GENERAL INFORMATION:
APPLICANT: Schmidt, Robert R.
APPLICANT: Miller, Philip
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES
RELATING TO THE - AND -SUBUNITS OF GLUTAMATE
TITLE OF INVENTION: DEHYDROGENASES AND METHODS OF USE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSER: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/541,033A
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI155
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1506 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 4..1464
US-08-541-033A-23

Query Match 3.7%; Score 32.8; DB 2; Length 1506;
Best Local Similarity 45.4%; Pred. No. 3.7;
Matches 118; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

OY 285 ggaagacatcgaaacatgttcaaaagcggtacacgctggaacacgtgacagcagctt 344
DB 1005 GTACAGAGCGGACACCGCGGTGATGTGGCGACCGCGCAAGCTTTGGAGCTG 1064
OY 345 ggaacaaggcggaaggctgtgttcaacagcgccgacatcgagctagattggcg 404
DB 1065 CCAGGTGACATCGCTTCCCTCGCGCACCCAGAACGAGATGATGACGACGCGCA 1124
OY 405 aegctacatcagccagcagcttcggttccactgacccgcatgtacaagccgcgaat 464
DB 1125 GGTGCTGATCAGACGCGCTGCCAGTACGTGTGAGGCGCCACATGCTCCACCAA 1184
OY 465 caaagcgatagaacaaatcatgtagcgcgaggtgctggaagcaaaacgcgcgc 524
DB 1185 CGAGGCGATCCACAACTATACAAAGCGCGCATCATCTACTGCCCGCAAGCGCCAA 1244
OY 525 caccggcatatacaagggtca 544
DB 1245 CGCGGCGCGCGTGGCGGTCA 1264

RESULT 29
US-08-828-451-23
Sequence 23, Application US/08828451
Patent No. 5985634
GENERAL INFORMATION:
APPLICANT: Schmidt, Robert R.
APPLICANT: Miller, Philip
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES
TITLE OF INVENTION: RELATING TO THE - AND -SUBUNITS OF GLUTAMATE
TITLE OF INVENTION: DEHYDROGENASES AND METHODS OF USE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/828,451
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/541,033
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UFI155
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 1506 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 4..1464
US-08-828-451-23

Query Match 3.7%; Score 32.8; DB 2; Length 1506;
Best Local Similarity 45.4%; Pred. No. 3.7;
Matches 118; conservative 0; Mismatches 142; Indels 0; Gaps 0;

OY 285 ggaagacatcgaaacatgttcaaaagcggtacacgctggaacacgtgacagcagctt 344
DB 1005 GTACAGAGCGGACACCGCGGTGATGTGGCGACCGCGCAAGCTTTGGAGCTG 1064
OY 345 ggaacaaggcggaaggctgtgttcaacagcgccgacatcgagctagattggcg 404
DB 1065 CCAGGTGACATCGCTTCCCTCGCGCACCCAGAACGAGATGATGACGACGCGCA 1124
OY 405 aegctacatcagccagcagcttcggttccactgacccgcatgtacaagccgcgaat 464
DB 1125 GGTGCTGATCAGACGCGCTGCCAGTACGTGTGAGGCGCCACATGCTCCACCAA 1184
OY 465 caaagcgatagaacaaatcatgtagcgcgaggtgctggaagcaaaacgcgcgc 524
DB 1185 CGAGGCGATCCACAACTATACAAAGCGCGCATCATCTACTGCCCGCAAGCGCCAA 1244
OY 525 caccggcatatacaagggtca 544
DB 1245 CGCGGCGCGCGTGGCGGTCA 1264

RESULT 30
US-08-541-033A-7
Sequence 7, Application US/08541033A
Patent No. 587941
GENERAL INFORMATION:
APPLICANT: Schmidt, Robert R.
APPLICANT: Miller, Philip
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES
TITLE OF INVENTION: RELATING TO THE - AND -SUBUNITS OF GLUTAMATE
TITLE OF INVENTION: DEHYDROGENASES AND METHODS OF USE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
COMPUTER READABLE FORM:

ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF155
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2099 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 33..1568
US-08-541-033A-3

Query Match 3.7%; Score 32.8; DB 2; Length 2099;
Best Local Similarity 45.4%; Pred. No. 4.3;
Matches 118; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 285 ggaagacatcgaaacatgttcaaacggtacacgctgggaacacgtgcagcagctt 344
DB 1109 GTACAAAGGCGACACCGCGTGTATGTGGGAGCGCGCAAGCTTGAGAGCTG 1168
QY 345 ggaagagggggaaggggtgtgtttcatcacgcgcacatcgagcgtactgtgggg 404
DB 1169 CCAGGTGACATGCGCTTCCCTCGCCACCCAGAACGATGATGACGACGCGCGA 1228
QY 405 agctcatcacgcagcagcttcggttccacctgaccgcacatgtacagccgcgaat 464
DB 1229 GCGTGTGATCAACGCGCTGCCAGTACGTGTGGAGGCGCCCAACATGCGCTCCACCA 1288
QY 465 caaagcagatagaacaaatcatgtcagggcgaggggtgctggcggaagcaaacgcgc 524
DB 1289 CGAGGCGCATCCACAACTACACAAAGCGCGCATCTACTGCCCCGCAAGGCGCCAA 1348
QY 525 caacggcatacaaggggtca 544
DB 1349 CGCCGCGCGCGTGGCGGTCA 1368

RESULT 35

US-08-828-451-3
Sequence 3, Application US/08828451
Patent No. 5985634
GENERAL INFORMATION:
APPLICANT: Schmidt, Robert R.
APPLICANT: Miller, Philip
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES
TITLE OF INVENTION: RELATING TO THE - AND -SUBUNITS OF GLUTAMATE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Saliwanhik & Saliwanhik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/828,451
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/541,033
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.
REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF155
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2099 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 33..1568
US-08-828-451-3

Query Match 3.7%; Score 32.8; DB 2; Length 2099;
Best Local Similarity 45.4%; Pred. No. 4.3;
Matches 118; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 285 ggaagacatcgaaacatgttcaaacggtacacgctgggaacacgtgcagcagctt 344
DB 1109 GTACAAAGGCGACACCGCGTGTATGTGGGAGCGCGCAAGCTTGAGAGCTG 1168
QY 345 ggaagagggggaaggggtgtgtttcatcacgcgcacatcgagcgtactgtgggg 404
DB 1169 CCAGGTGACATGCGCTTCCCTCGCCACCCAGAACGATGATGACGACGCGCGA 1228
QY 405 agctcatcacgcagcagcttcggttccacctgaccgcacatgtacagccgcgaat 464
DB 1229 GCGTGTGATCAACGCGCTGCCAGTACGTGTGGAGGCGCCCAACATGCGCTCCACCA 1288
QY 465 caaagcagatagaacaaatcatgtcagggcgaggggtgctggcggaagcaaacgcgc 524
DB 1289 CGAGGCGCATCCACAACTACACAAAGCGCGCATCTACTGCCCCGCAAGGCGCCAA 1348
QY 525 caacggcatacaaggggtca 544
DB 1349 CGCCGCGCGCGTGGCGGTCA 1368

RESULT 36

US-08-541-033A-18
Sequence 18, Application US/08541033A
Patent No. 5879941
GENERAL INFORMATION:
APPLICANT: Schmidt, Robert R.
APPLICANT: Miller, Philip
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES
TITLE OF INVENTION: RELATING TO THE - AND -SUBUNITS OF GLUTAMATE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Saliwanhik & Saliwanhik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/541,033A
FILING DATE:

TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2140 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 33..1610
US-08-541-033A-1

Query Match 3.7%; Score 32.8; DB 2; Length 2140;
Best Local Similarity 45.4%; Pred. No. 4.4; Mismatches 142; Indels 0; Gaps 0;
Matches 118; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 285 ggaagacatcgaaacaatgttcaaaagcgttacacgctggaacaacgtgcagcagctt 344
DB 1151 gTCAAGAGCGACACCGCGTGTATGTGGGGAGACCGCCGCAAGCCTTGGAGCTGACTG 1210
QY 345 ggaagaggggaagggctgtgttcaacagcgccacacatcgcaatgattgggcgg 404
DB 1211 CCAAGTGGACATCGCTTCCCTGCGCCACCAAGAGATCGATGAGCAGACGCCGA 1270
QY 405 acgctacatcagcagcagcttcgcttccacctgaccgcacatgtaaacgcgcgaat 464
DB 1271 GCTGTGATCAAGACACCGCGTGTGAGTGGAGGGGCCCAACATGCCCTCCACCAA 1330
QY 465 caaagcgatagaaataatctgcagcgcggaaggtgcgcggaaggaacacgcgc 524
DB 1331 CGAGGCGATCCACAAGTACAAAGCGCGCATCTACTGCCCCGCAAGCGGCCAA 1390
QY 525 caccgcatacaaggggtca 544
DB 1391 CCGCGCGCGCGTGGCGGTCA 1410

RESULT 39
US-08-828-451-1
Sequence 1, Application US/08828451
Patent No. 5985634
GENERAL INFORMATION:
APPLICANT: Schmidt, Robert R.
APPLICANT: Miller, Philip
TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES
TITLE OF INVENTION: RELATING TO THE - AND -SUBUNITS OF GLUTAMATE
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/828,451
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/541,033
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Whitlock, Ted W.

REGISTRATION NUMBER: 36,965
REFERENCE/DOCKET NUMBER: UF155
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2140 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 33..1610
US-08-828-451-1

Query Match 3.7%; Score 32.8; DB 2; Length 2140;
Best Local Similarity 45.4%; Pred. No. 4.4; Mismatches 142; Indels 0; Gaps 0;
Matches 118; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 285 ggaagacatcgaaacaatgttcaaaagcgttacacgctggaacaacgtgcagcagctt 344
DB 1151 gTCAAGAGCGACACCGCGTGTATGTGGGGAGACCGCCGCAAGCCTTGGAGCTGACTG 1210
QY 345 ggaagaggggaagggctgtgttcaacagcgccacacatcgcaatgattgggcgg 404
DB 1211 CCAAGTGGACATCGCTTCCCTGCGCCACCAAGAGATCGATGAGCAGACGCCGA 1270
QY 405 acgctacatcagcagcagcttcgcttccacctgaccgcacatgtaaacgcgcgaat 464
DB 1271 GCTGTGATCAAGACACCGCGTGTGAGTGGAGGGGCCCAACATGCCCTCCACCAA 1330
QY 465 caaagcgatagaaataatctgcagcgcggaaggtgcgcggaaggaacacgcgc 524
DB 1331 CGAGGCGATCCACAAGTACAAAGCGCGCATCTACTGCCCCGCAAGCGGCCAA 1390
QY 525 caccgcatacaaggggtca 544
DB 1391 CCGCGCGCGCGTGGCGGTCA 1410

RESULT 40
US-09-105-537-1
Sequence 1, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,438051
CURRENT APPLICATION NUMBER: US/09/105,537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 15872
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-105-537-1

Query Match 3.7%; Score 32.8; DB 4; Length 15872;
Best Local Similarity 56.5%; Pred. No. 11;
Matches 61; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 520 gccgcacgcgcatacagaaggggtcaaaacaatcatcaagccctgcgcgagcgagga 579
DB 9548 gccgcgcgcggtgtcgcgcggtcatcaagatgtctcagcgcgctgcgcagcgttgc 9607

QY 580 accatcatctgcccagaccagctcccttctcgaggaagcgcgagc 627
11 11 1111 11 1111 11 11111 11 11
Db 9608 ccccgaccctccagcgagccgagcccgatgcccactggagc 9655

RESULT 41

US-08-804-227C-1
; Sequence 1, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: Dehoff, Bradley S.
; APPLICANT: Kustoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43280 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 816..14234
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14351..19945
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20010..31199
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31232..36067
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36249..41774
; US-08-804-227C-1

Query Match 3.7%; Score 32.8; DB 2; Length 43280;
Best Local Similarity 56.5%; Pred. No. 17;

Matches 61; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 520 gcgcaccacggcatacaagggtcaacaacatcaagccctgcgcgagcgagga 579
11 11 1111 11 1111 11 11111 1111 1111
Db 37410 gccgcggcggtgctgcggcgctcattcaagatgctccagagcgctgcgcttgctg 37469
11 11 1111 11 1111 11 11111 11 1111
QY 580 accatcatctgcccagaccagctcccttctcgaggaagcgcgagc 627
11 11 1111 11 1111 11 11111 11 11

Db 37470 ccccgaccctccagcgcgagccgagcccgatgcccactggagc 37517

RESULT 42

US-08-997-080-179
; Sequence 179, Application US/08997080
; Patent No. 5968524
; GENERAL INFORMATION:
; APPLICANT: NATSON, JAMES D.
; APPLICANT: TAN, PAUL L.J.
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,080
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 179:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 520 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-997-080-179

Query Match 3.6%; Score 32.6; DB 2; Length 520;
Best Local Similarity 49.7%; Pred. No. 2.7;
Matches 83; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 303 gtltcaagcggtacagcgctgggaacacgtcagcagcgcttggacaagggcgaggt 362
11 1111 11 11 1111 11 11111 11 111111 11
Db 312 gcttaccgagagcgccgcccgcgacctgtgtcaagagccgaaggcgagagccg 371
11 11 1111 11 1111 11 11111 11 111111
QY 363 gctgtcatcagcgccgacatcgagctacgatttggcgagcgctacatcagcagca 422
11 11 1111 11 1111 11 1111 11 1111
Db 372 CAAGGTGTGGTGGCAGACATCCGTGCGACGATATGAACACCTTTCGATGCGACGGGT 431
11 11 1111 11 1111 11 11 11 1111
QY 423 gcttcggttccacctgagcgcgcatgtacaagcgccggaatcaaga 469
11 11 1111 11 1111 11 11 11 1111
Db 432 AGGCTGCCGAGCCGCCCGCTCTGTAGAACGACAGAGAGATCGCAG 478
11 11 1111 11 1111 11 11 11 1111

RESULT 43

US-08-997-362-179
; Sequence 179, Application US/08997362
; Patent No. 5985287
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth


```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 179
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Mycobacterium vaccae
US-09-324-542-179
```

```
Query Match          3.6%; Score 32.6; DB 4; Length 520;
Best Local Similarity 49.7%; Pred. No. 2.7;
Matches 83; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
```

```
OY 303 gttcaaacggtacacggtctgggaacacgtgacgacgcttggacaaggcgaaaggt 362
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 312 gttcacgagagagcgcccgccgacctgtgtcaagcaggccaaggcgagagacgc 371
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 363 gctgtcatcacgacgacatcggaactaagattggcggaagcgtacatcaagccagca 422
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 372 caaggtgtcggtgcaacacatccgtcgcaacgatatgaacacattcgcatcgacccgt 431
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 423 gttccgttccacctgacgacgcatgtacaagcgccgaaatacaag 469
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 432 acggctgacgacgacccgctcgtctagaagcagagagatcgca 478
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: July 1, 2002, 04:28:32
Job time: 113437 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2002, 01:19:33 ; Search time 415.2 Seconds
(without alignments)
530.668 Million cell updates/sec

Title: US-09-303-518D-569

Perfect score: 897
Sequence: 1 atggttcggttaacatcag.....accgctacaacatgcgtaa 897

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 12816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued Patents_NA:*

- 1: /cgn2.6/ptodata/1/lna/5A.COMB.seq:*
- 2: /cgn2.6/ptodata/1/lna/5B.COMB.seq:*
- 3: /cgn2.6/ptodata/1/lna/6A.COMB.seq:*
- 4: /cgn2.6/ptodata/1/lna/6B.COMB.seq:*
- 5: /cgn2.6/ptodata/1/lna/PCFUS.COMB.seq:*
- 6: /cgn2.6/ptodata/1/lna/Backfile1.seq:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	36	4.0	7218	1	US-08-232-463-14
C 2	35.6	4.0	4411529	4	US-09-103-840A-1
C 3	33.8	3.8	1838	3	US-08-470-260-7
C 4	33.8	3.8	1838	3	US-08-471-491-7
C 5	33.8	3.8	1838	3	US-08-466-662-7
C 6	32.2	3.6	289	4	US-09-007-005-17
C 7	32.2	3.6	289	4	US-09-244-796-17
C 8	32.2	3.6	7218	1	US-08-232-463-14
C 9	32.2	3.6	4403765	2	US-09-103-840A-2
C 10	32	3.6	1619	4	US-08-991-946A-2
C 11	31.8	3.5	1483	3	US-09-262-749-1
C 12	31.8	3.5	1553	3	US-09-022-669-1
C 13	31.6	3.5	289	4	US-09-007-005-17
C 14	31.6	3.5	289	4	US-09-244-796-17
C 15	31.6	3.5	496	4	US-08-905-223-23
C 16	31.6	3.5	496	4	US-09-247-155-23
C 17	31.6	3.5	1119	4	US-08-987-943-2
C 18	31.4	3.5	1755	4	US-08-987-943-1
C 19	31.4	3.5	5247	1	US-08-920-812-15
C 20	31.4	3.5	5247	1	US-08-920-827-15
C 21	31.4	3.5	5247	1	US-08-921-177-15
C 22	31.4	3.5	5247	1	US-08-362-577C-15
C 23	31.4	3.5	5247	2	US-08-920-828-15
C 24	31.2	3.5	11958	4	US-09-134-246-8
C 25	31	3.5	480	3	US-08-814-052-15
C 26	31	3.5	1082	2	US-08-715-252-1
C 27	31	3.5	1082	2	US-08-453-051-3

C 28	31	3.5	1130	1	US-08-406-948A-5	Sequence 5, Appl1
C 29	30	3.3	3358	4	US-09-248-571-2	Sequence 2, Appl1
C 30	30	3.3	6816	4	US-09-404-650-1	Sequence 1, Appl1
C 31	30	3.3	6855	4	US-09-404-650-3	Sequence 3, Appl1
C 32	29.8	3.3	2478	1	US-08-481-626-1	Sequence 1, Appl1
C 33	29.8	3.3	4020	4	US-09-050-159-130	Sequence 130, App
C 34	29.8	3.3	4024	4	US-09-162-484-18	Sequence 18, Appl1
C 35	29.6	3.3	2256	4	US-07-794-393-1	Sequence 1, Appl1
C 36	29.6	3.3	2256	1	US-08-001-711-1	Sequence 1, Appl1
C 37	29.4	3.3	11219	1	US-07-642-734C-1	Sequence 1, Appl1
C 38	29.4	3.3	11219	3	US-08-439-009A-1	Sequence 196, App
C 39	29.2	3.3	342	4	US-08-905-223-196	Sequence 3, Appl1
C 40	29.2	3.3	1432	4	US-07-914-282D-3	Sequence 3, Appl1
C 41	29.2	3.3	1432	1	US-08-276-887A-3	Sequence 3, Appl1
C 42	29.2	3.3	1432	5	PCT-US93-02460-3	Sequence 3, Appl1
C 43	29.2	3.3	1683	1	US-07-945-283-3	Sequence 4, Appl1
C 44	29.2	3.3	1766	4	US-09-321-981-4	Sequence 3, Appl1
C 45	29.2	3.3	1782	2	US-08-714-168-3	Sequence 3, Appl1
C 46	29.2	3.3	1782	4	US-09-320-721A-3	Sequence 3, Appl1
C 47	29.2	3.3	1831	6	5215881-1	Patent No. 5215881
C 48	29.2	3.3	1831	3	5215881-3	Patent No. 5215881
C 49	29.2	3.3	2223	1	US-08-317-522A-8	Sequence 8, Appl1
C 50	29.2	3.3	2223	2	US-08-439-818A-8	Sequence 8, Appl1
C 51	29.2	3.3	2223	2	US-08-751-965-8	Sequence 8, Appl1
C 52	29.2	3.3	2223	2	US-08-738-973-8	Sequence 8, Appl1
C 53	29.2	3.3	2223	2	US-08-728-626-8	Sequence 8, Appl1
C 54	29.2	3.3	2223	2	US-08-808-599A-8	Sequence 8, Appl1
C 55	29	3.2	1095	1	US-08-385-186-8	Sequence 12, Appl1
C 56	29	3.2	1425	1	US-08-385-186-12	Sequence 5, Appl1
C 57	29	3.2	1639	1	US-08-334-698-5	Sequence 5, Appl1
C 58	29	3.2	1639	1	US-08-228-932-5	Sequence 5, Appl1
C 59	29	3.2	1639	1	US-08-468-939-5	Sequence 5, Appl1
C 60	29	3.2	1639	2	US-08-406-855A-5	Sequence 5, Appl1
C 61	29	3.2	1639	2	US-08-722-190-5	Sequence 5, Appl1
C 62	29	3.2	1639	3	US-08-244-354-5	Sequence 5, Appl1
C 63	29	3.2	1639	3	US-09-206-899-5	Sequence 5, Appl1
C 64	29	3.2	1639	5	PCT-US95-04203-5	Sequence 5, Appl1
C 65	29	3.2	7122	4	US-09-318-448-2	Sequence 2, Appl1
C 66	29	3.2	11492	4	US-08-991-840A-1	Sequence 1, Appl1
C 67	29	3.2	4411529	4	US-09-103-840A-1	Sequence 1, Appl1
C 68	28.8	3.2	4800	3	US-08-941-445A-4	Sequence 4, Appl1
C 69	28.8	3.2	6226	1	US-08-542-363-1	Sequence 1, Appl1
C 70	28.8	3.2	6226	4	US-09-100-089-1	Sequence 1, Appl1
C 71	28.8	3.2	35081	2	US-08-752-760A-1	Sequence 1, Appl1
C 72	28.6	3.2	4403765	4	US-09-103-840A-2	Sequence 2, Appl1
C 73	28.6	3.2	1424	1	US-08-403-634-3	Sequence 3, Appl1
C 74	28.6	3.2	1424	1	US-08-913-441B-3	Sequence 3, Appl1
C 75	28.6	3.2	1611	5	PCT-US93-08743-3	Sequence 3, Appl1
C 76	28.6	3.2	5925	3	US-08-470-260-4	Sequence 4, Appl1
C 77	28.6	3.2	5925	3	US-08-471-491-4	Sequence 4, Appl1
C 78	28.6	3.2	5925	3	US-08-466-662-4	Sequence 4, Appl1
C 79	28.6	3.2	10299	2	US-08-477-451-1	Sequence 1, Appl1
C 80	28.6	3.2	10299	2	US-08-477-451-5	Sequence 5, Appl1
C 81	28.6	3.2	19932	2	US-08-477-451-25	Sequence 25, Appl1
C 82	28.4	3.2	434	4	US-09-060-756-67	Sequence 67, Appl1
C 83	28.4	3.2	476	4	US-08-050-259B-19	Sequence 19, Appl1
C 84	28.4	3.2	1278	4	US-09-613-182-14	Sequence 14, Appl1
C 85	28.4	3.2	2846	4	US-09-613-182-5	Sequence 5, Appl1
C 86	28.4	3.2	2963	4	US-09-232-200-60	Sequence 60, Appl1
C 87	28.4	3.2	2963	4	US-09-232-197-60	Sequence 60, Appl1
C 88	28.4	3.2	2963	4	US-09-232-201-60	Sequence 60, Appl1
C 89	28.4	3.2	4529	1	US-08-565-386-1	Sequence 1, Appl1
C 90	28.2	3.1	3416	2	US-08-701-240-3	Sequence 1, Appl1
C 91	28.2	3.1	14683	2	US-09-138-236-3	Sequence 3, Appl1
C 92	28.2	3.1	14683	2	US-08-819-866-1	Sequence 1, Appl1
C 93	28.2	3.1	1393	2	US-09-023-715-1	Sequence 1, Appl1
C 94	28	3.1	1393	3	US-09-011-197-3	Sequence 3, Appl1
C 95	28	3.1	1359	4	US-08-582-740-69	Sequence 69, Appl1
C 96	28	3.1	1539	4	US-09-109-879-69	Sequence 69, Appl1
C 97	28	3.1	1624	3	US-08-582-740-67	Sequence 67, Appl1
C 98	28	3.1	1624	4	US-09-109-879-67	Sequence 67, Appl1
C 99	28	3.1	2361	4	US-09-011-197-1	Sequence 1, Appl1
C 100	28	3.1	3636	3	US-09-074-579-2	Sequence 2, Appl1


```
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 496 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Cancerous prostate
; FEATURE:
; NAME/KEY: other
; LOCATION: 149..331
; IDENTIFICATION METHOD: blastn
; OTHER INFORMATION: identity 98
; OTHER INFORMATION: region 1..183
; OTHER INFORMATION: id AA397994
; OTHER INFORMATION: est
; FEATURE:
; NAME/KEY: other
; LOCATION: 328..485
; IDENTIFICATION METHOD: blastn
; OTHER INFORMATION: identity 96
; OTHER INFORMATION: region 179..336
; OTHER INFORMATION: id AA397994
; OTHER INFORMATION: est
; FEATURE:
; NAME/KEY: complement(182..496)
; LOCATION: 196..240
; IDENTIFICATION METHOD: blastn
; OTHER INFORMATION: identity 97
; OTHER INFORMATION: region 14..328
; OTHER INFORMATION: id AA39680
; OTHER INFORMATION: est
; FEATURE:
; NAME/KEY: sig-peptide
; LOCATION: 196..240
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 5.5
; OTHER INFORMATION: seq ILSTVATLTFAXA/LD
; US-08-905-223-23

Query Match
Best Local Similarity 50.7%; Score 31.6; DB 4; Length 496;
Matches 73; Conservative 1; Mismatches 70; Indels 0; Gaps 0;

QY 591 gccgcacacgctcccccctcaagaagcgaggcgatgtggtgattcttcg 650
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 120 gccggagatagagaccacgcctcagagatgttttcttcggacgtatcga 179
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 651 caaacctgctataccatgacgtgcgcaaaattggcacacgtcaaggcgtgaaac 710
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 180 gccacacagacacacacgtggtggtctgctacagtcacgcttaacattgccaragc 239
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 711 cctgtttctgtcgcaagcct 734
    ||||| ||||| | | | |
Db 240 cctggacgctgcgaagaatgcat 263
    ||||| ||||| | | | |

RESULT 16
US-09-247-155-23
; Sequence 23, Application US/09247155A
; Patent No. 6312922
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bouguetieret, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET 021A
; CURRENT APPLICATION NUMBER: US/09/247,155A
; CURRENT FILING DATE: 1999-02-09
```

```
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-04-13
; EARLIER APPLICATION NUMBER: 60/096,116
; EARLIER FILING DATE: 1998-08-10
; EARLIER APPLICATION NUMBER: 60/099,273
; EARLIER FILING DATE: 1998-10-04
; NUMBER OF SEQ ID NOS: 182
; SOFTWARE: Patent.pm
; SEQ ID NO 23
; LENGTH: 496
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 149..331
; OTHER INFORMATION: blastn
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 328..485
; OTHER INFORMATION: blastn
; FEATURE:
; NAME/KEY: complement(182..496)
; LOCATION: 196..240
; OTHER INFORMATION: Von Heijne matrix
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 101
; OTHER INFORMATION: n-a, g, c or t
; US-09-247-155-23

Query Match
Best Local Similarity 50.7%; Score 31.6; DB 4; Length 496;
Matches 73; Conservative 1; Mismatches 70; Indels 0; Gaps 0;

QY 591 gccgcacacgctcccccctcaagaagcgaggcgatgtggtgattcttcg 650
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 120 gccggagatagagaccacgcctcagagatgttttcttcggacgtatcga 179
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 651 caaacctgctataccatgacgtgcgcaaaattggcacacgtcaaggcgtgaaac 710
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 180 gccacacagacacacacgtggtggtctgctacagtcacgcttaacattgccaragc 239
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 711 cctgtttctgtcgcaagcct 734
    ||||| ||||| | | | |
Db 240 cctggacgctgcgaagaatgcat 263
    ||||| ||||| | | | |

RESULT 17
US-08-987-943-2
; Sequence 2, Application US/08987943
; Patent No. 6210967
; GENERAL INFORMATION:
; APPLICANT: Bard, Jonathan A.
; TITLE OF INVENTION: DNA Encoding A Mammalian LPA Receptor and Uses Thereof
; FILE REFERENCE: 55183
; CURRENT APPLICATION NUMBER: US/08/987,943
; CURRENT FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0 - beta
; SEQ ID NO 2
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-08-987-943-2
```

Query Match 3.5%; Score 31.4; DB 4; Length 1119;
Best Local Similarity 61.7%; Pred. No. 3.1;
Matches 50; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

28 ccccttgcaaccgcatcctgttgcacgcctctcaaatgctctccg 87
11 11 1111 111111 11 1111 111111 1 1 1 11
Db 187 ccgcatatgcctgcggcacaacctgtgtgcgcacacctgctctccgacatcatg 246

QY 88 ctgcgccttctctgtctgcac 108
11 11 11 11 11 11
Db 247 ctgcctctgtcatgccc 267

RESULT 18
US-08-987-943-1
; Sequence 1, Application US/08987943
; Patent No. 6210967
; GENERAL INFORMATION:
; APPLICANT: Bard, Jonathan A.
; TITLE OF INVENTION: DNA Encoding A Mammalian LPA Receptor and Uses Thereof
; FILE REFERENCE: 55183
; CURRENT APPLICATION NUMBER: US/08/987,943
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 1
; LENGTH: 1755
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-987-943-1

Query Match 3.5%; Score 31.4; DB 4; Length 1755;
Best Local Similarity 61.7%; Pred. No. 4;
Matches 50; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

28 ccccttgcaaccgcatcctgttgcacgcctctcaaatgctctccg 87
11 11 1111 111111 11 1111 111111 1 1 1 11
Db 536 ccgcatatgcctgcggcacaacctgtgtgcgcacacctgctctccgacatcatg 595

QY 88 ctgcgccttctctgtctgcac 108
11 11 11 11 11 11
Db 596 ctgcctctgtcatgccc 616

RESULT 19
US-08-920-812-15/c
; Sequence 15, Application US/08920812
; Patent No. 5763188
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,812
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 5247 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Pseudomonas aeruginosa
STRAIN: Clinical Isolate P2-17
US-08-920-812-15

Query Match 3.5%; Score 31.4; DB 1; Length 5247;
Best Local Similarity 64.4%; Pred. No. 7.2;
Matches 47; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

42 cggcatatgcctgcggcacaacctgtgtgcacacctgctctccgacatcatg 101
11 11 1111 11 11 1111111111 11 1111 1111 1
Db 698 CGCATATGCTTCGGCTGCACCGCCCTGCTCATGACCATCTCTGTGCGCTGTTCCT 639

QY 102 tctgcacacgctg 114
11 11 11 11 11
Db 638 CGCGCCACGCTG 626

RESULT 20
US-08-920-827-15/c
; Sequence 15, Application US/08920827
; Patent No. 5770375
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,827
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448

TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 5247 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Pseudomonas aeruginosa
STRAIN: Clinical Isolate P2-17
US-08-920-827-15

Query Match 3.5%; Score 31.4; DB 1; Length 5247;
Best Local Similarity 64.4%; Pred. No. 7.2;
Matches 47; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 42 cgcacatcacatctgtttaccgcgcctgtcacaatgacctctccctgtgcgcgttctctg 101
||||| ||| | ||||| ||||| ||| ||| | ||| |
Db 698 CGCATGTACTTCGGCTGGACCGCCTGCTCATGACCATCTCTGTTGGCGCTTCTCT 639

QY 102 tctgcaacagctg 114
|||||
Db 638 CGGCGCCACGCTG 626

RESULT 21

US-08-921-177-15/c
Sequence 15, Application US/08921177
Patent No. 5798211

GENERAL INFORMATION:

APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921.177
FILING DATE: 29-AUG-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 5247 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:

ORGANISM: Pseudomonas aeruginosa
STRAIN: Clinical Isolate P2-17
US-08-921-177-15

Query Match 3.5%; Score 31.4; DB 1; Length 5247;
Best Local Similarity 64.4%; Pred. No. 7.2;
Matches 47; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 42 cgcacatcacatctgtttaccgcgcctgtcacaatgacctctccctgtgcgcgttctctg 101
||||| ||| | ||||| ||||| ||| ||| | ||| |
Db 698 CGCATGTACTTCGGCTGGACCGCCTGCTCATGACCATCTCTGTTGGCGCTTCTCT 639

QY 102 tctgcaacagctg 114
|||||
Db 638 CGGCGCCACGCTG 626

RESULT 22

US-08-362-577C-15/c
Sequence 15, Application US/08362577C
Patent No. 5807673

GENERAL INFORMATION:

APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,577C
FILING DATE: 27-MAR-1995

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:
LENGTH: 5247 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Pseudomonas aeruginosa
STRAIN: Clinical Isolate P2-17
US-08-362-577C-15

Query Match 3.5%; Score 31.4; DB 1; Length 5247;
Best Local Similarity 64.4%; Pred. No. 7.2;
Matches 47; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 42 cgcacatcacatctgtttaccgcgcctgtcacaatgacctctccctgtgcgcgttctctg 101
||||| ||| | ||||| ||||| ||| ||| | ||| |
Db 698 CGCATGTACTTCGGCTGGACCGCCTGCTCATGACCATCTCTGTTGGCGCTTCTCT 639


```

1 STRANDEDNESS: single
2 TOPOLOGY: linear
3 MOLECULE TYPE: DNA (genomic)
4 FEATURE:
5   NAME/KEY: misc_feature
6   LOCATION: 1..6
7   OTHER INFORMATION: /function= "BamHI linker"
8 FEATURE:
9   NAME/KEY: promoter
10  LOCATION: 7..432
11  OTHER INFORMATION: /standard_name= "CupI promoter"
12 FEATURE:
13   NAME/KEY: misc_feature
14   LOCATION: 433..441
15   OTHER INFORMATION: /function= "EcoRI linker"
16 FEATURE:
17   NAME/KEY: sig_peptide
18   LOCATION: 442..492
19   OTHER INFORMATION: /standard_name= "PH05 signal"
20   OTHER INFORMATION: sequence"
21 FEATURE:
22   NAME/KEY: mat_peptide
23   LOCATION: 493..690
24   OTHER INFORMATION: /product= "desulphatohirudin"
25   OTHER INFORMATION: /standard_name= "Hv1"
26 FEATURE:
27   NAME/KEY: terminator
28   LOCATION: 691..1068
29   OTHER INFORMATION: /standard_name= "PH05 terminator"
30 FEATURE:
31   NAME/KEY: misc_feature
32   LOCATION: 1069..1082
33   OTHER INFORMATION: /function= "SalI linker"
34 FEATURE:
35   NAME/KEY: CDS
36   LOCATION: 442..690
37   OTHER INFORMATION: /product= "primary transcript"
38 US-08-453-051-3
39
40 Query Match 3.5%; Score 31; DB 2; Length 1082;
41 Best Local Similarity 57.9%; Pred. No. 4.1;
42 Matches 55; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
43
44 Oy 656 caaagcgctgaaaccctgtttctgctggaacgcctgcctgcggaagaaggttcga 755
45 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
46 Db 852 CATACTCGTAGAAGTCATTGATTTCACAAGAGAACCCCTGGACCAAGTGAACGTTTCA 793
47
48 Oy 756 ttgcacatccgccccgcgcccaagggaattgaacg 790
49 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
50 Db 792 TTGGAACAACAGCATCGTTAATGACGTATCTGCAGC 758
51
52 RESULT 28
53 US-08-406-948A-5/C
54 ; Sequence 5, Application US/08406948A
55 ; Patent No. 5661001
56 ; GENERAL INFORMATION:
57 ; APPLICANT: Grossenbacher, Hugh
58 ; APPLICANT: Jui Yoa, Chang
59 ; APPLICANT: Walter, M Ki
60 ; TITLE OF INVENTION: High Molecular Weight Hirudin
61 ; NUMBER OF SEQUENCES: 8
62 ; CORRESPONDENCE ADDRESS:
63 ; ADDRESSEE: CIBA-GEIGY Corporation
64 ; STREET: 520 White Plains Road
65 ; CITY: Tarrytown
66 ; STATE: NY
67 ; ZIP: 10591-9005
68 ; COMPUTER READABLE FORM:
69 ; MEDIUM TYPE: Floppy disk
70 ; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/406,948A
FILING DATE: March 29, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP94/02438
FILING DATE: 23 JUL 94
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 4-19636/A/PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-3318
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1130 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc.signal
LOCATION: 1..1130
OTHER INFORMATION: /function= "first 1130 nucleotides"
OTHER INFORMATION: of pDDB207/GAPFL-YHR (EP-A-340 170)."
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..276
OTHER INFORMATION: /product= "SalI - BamHI fragment of
OTHER INFORMATION: pBR322"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 277..284
OTHER INFORMATION: /function= "BamHI/BglII linker"
FEATURE:
NAME/KEY: promoter
LOCATION: 285..478
OTHER INFORMATION: /function= "GAPFL promoter"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 479..484
OTHER INFORMATION: /product= "EcoRI linker"
FEATURE:
NAME/KEY: misc.signal
LOCATION: 488..538
OTHER INFORMATION: /function= "PHO5 signal sequence"
FEATURE:
NAME/KEY: mat.peptide
LOCATION: 539..733
OTHER INFORMATION: /product= "Hlrudin HVI"
OTHER INFORMATION: /standard_name= "HVI"
FEATURE:
NAME/KEY: terminator
LOCATION: 737..1113
OTHER INFORMATION: /standard_name= "PHO5 terminator"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1117..1122
OTHER INFORMATION: /function= "M13 cloning site"
FEATURE:
NAME/KEY: misc.signal
LOCATION: 734..736
OTHER INFORMATION: /function= "translation terminator"
OTHER INFORMATION: /standard_name= "TAG"

```

Query Match	3.5%;	Score 31;	DB 1;	Length 1130;
Best Local Similarity	57.9%;	Pred. No. 4.2;		
Matches	55;	Conservative	0;	Mismatches 40;
				Indels 0;
				Gaps 0;


```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,626
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/656,183
; FILING DATE: 04-MAR-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 89-09062
; FILING DATE: 05-JUL-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 04958-0006-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2478 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-481-626-1
```

```
Query Match 3.3%; Score 29.8; DB 1; Length 2478;
Best Local Similarity 47.6%; Pred. No. 16;
Matches 88; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
```

```
OY 88 ctgcgcgttcctcgtctgcacacgctgggaaacgcgctcgcacatctgcgctttacctt 147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1825 CACCTTCATGCGGTGCGCCAGCGCTGCCGCTCTTGACCTGTGATGTCACACTT 1766
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 148 ttaagaaagacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 207
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1765 GTGCAGGGGGCCCTGTGGCCAGCTGCTGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1706
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 208 cccaacacggtlcaaacgcgcttttgcggaacgcgcaaaagcggttggaaactgcccc 267
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1705 GATGAAGCTGACAAAGTACGTGATGTAGACGACGCTAGAGAAGATGTGAAGTGGCCCC 1646
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 268 gcgctt 272
    |||
DB 1645 TGGGT 1641
```

```
RESULT 33
US-09-050-159-130/C
; Sequence 130, Application US/09050159A
; Patent No. 6197505
; GENERAL INFORMATION:
; APPLICANT: No. 6197505berg, Ielf T
; APPLICANT: Andersson, Maria K
; APPLICANT: Linstrom, Per H
; TITLE OF INVENTION: METHODS FOR ASSESSING CARDIOVASCULAR STATUS AND
; TITLE OF INVENTION: COMPOSITIONS FOR USE THEREOF
; FILE REFERENCE: 1248/1D042
; CURRENT APPLICATION NUMBER: US/09/050,159A
; CURRENT FILING DATE: 1998-03-27
; EARLIER APPLICATION NUMBER: 60/042,930
; EARLIER FILING DATE: 1987-04-03
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 130
; LENGTH: 4020
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; OTHER INFORMATION: Angiotensin I converting enzyme mRNA
US-09-050-159-130
```

```
Query Match 3.3%; Score 29.8; DB 4; Length 4020;
Best Local Similarity 47.6%; Pred. No. 20;
Matches 88; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
```

```
OY 88 ctgcgcgttcctcgtctgcacacgctgggaaacgcgctcgcacatctgcgctttacctt 147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3541 CACCTTCATGCGGTGCGCCAGCGCTGCCGCTCTTGACCTGTGATGTCACACTT 3482
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 148 ttaagaaagacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 207
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3481 GTGCAGGGGGCCCTGTGGCCAGCTGCTGGCAGACGTGCTGCTGCTGCTGCTGCTGCTGCTG 3422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 208 cccaacacggtlcaaacgcgcttttgcggaacgcgcaaaagcggttggaaactgcccc 267
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3421 GATGAAGCTGACAAAGTACGTGATGTAGACGACGCTAGAGAAGATGTGAAGTGGCCCC 3362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 268 gcgctt 272
    |||
DB 3361 TGGGT 3357
```

```
RESULT 34
US-09-162-484-18/C
; Sequence 18, Application US/09162484
; Patent No. 6248724
; GENERAL INFORMATION:
; APPLICANT: Phillips, M. Ian
; APPLICANT: Mohuczy, Dagnara
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDE COMPOSITIONS TARGETED TO
; TITLE OF INVENTION: CONVERTING ENZYME MRNA AND METHODS OF USE
; FILE REFERENCE: UFLA:087/UFLA087P
; CURRENT APPLICATION NUMBER: US/09/162,484
; CURRENT FILING DATE: 1998-09-25
; EARLIER APPLICATION NUMBER: 60/059,661
; EARLIER FILING DATE: 1997-09-25
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 18
; LENGTH: 4024
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-162-484-18
```

```
Query Match 3.3%; Score 29.8; DB 4; Length 4024;
Best Local Similarity 47.6%; Pred. No. 20;
Matches 88; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
```

```
OY 88 ctgcgcgttcctcgtctgcacacgctgggaaacgcgctcgcacatctgcgctttacctt 147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3541 CACCTTCATGCGGTGCGCCAGCGCTGCCGCTCTTGACCTGTGATGTCACACTT 3482
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 148 ttaagaaagacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 207
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3481 GTGCAGGGGGCCCTGTGGCCAGCTGCTGGCAGACGTGCTGCTGCTGCTGCTGCTGCTGCTG 3422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 208 cccaacacggtlcaaacgcgcttttgcggaacgcgcaaaagcggttggaaactgcccc 267
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3421 GATGAAGCTGACAAAGTACGTGATGTAGACGACGCTAGAGAAGATGTGAAGTGGCCCC 3362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 268 gcgctt 272
    |||
DB 3361 TGGGT 3357
```

```
RESULT 35
US-07-794-393-1/C
; Sequence 1, Application US/07794393
; Patent No. 5236844
```


NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Steven F. Weinstein
 STREET: Abbott Laboratories D377/AP6D-2 One Abbott
 STREET: Park Rd
 CITY: Abbott Park
 STATE: IL
 COUNTRY: US
 ZIP: 60064-3500
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/439,009A
 FILING DATE: 11-MAY-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Casuto, Diane
 REGISTRATION NUMBER: 40,943
 REFERENCE/DOCKET NUMBER: 4952-US.D1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 847-938-3137
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 11219 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: *Saccharopolyspora erythraea*
 STRAIN: NRRL 2338
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: 744..6659
 OTHER INFORMATION: /function= "APPROXIMATE SPAN OF
 OTHER INFORMATION: MODULE 1"
 OTHER INFORMATION: /label= FUNCTION
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 744..11219
 OTHER INFORMATION: /function= "gene= "eryA"
 OTHER INFORMATION: /product= "ORF1 encoding modules 1 & 2 for
 OTHER INFORMATION: 6-deoxyerythronolide B"
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: 744..1868
 OTHER INFORMATION: /function= "approximate span of
 OTHER INFORMATION: acyltransferase domain 1 of module 1"
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: 1998..2198
 OTHER INFORMATION: /function= "approximate span of
 OTHER INFORMATION: acyl carrier domain 1 of module 1"
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: 2250..3626
 OTHER INFORMATION: /function= "approximate span of
 OTHER INFORMATION: beta-ketoacylACP synthase domain/module1"
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: 3831..4811
 OTHER INFORMATION: /function= "approximate span of
 OTHER INFORMATION: acyltransferase 2 domain of module 1"
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: 5574..6125
 OTHER INFORMATION: /function= "approximate span of
 OTHER INFORMATION: beta-ketoreductase domain of module 1"

```

? FEATURE: misc_feature
? NAME/KEY: 6369_6626
? LOCATION: 6369..6626
? OTHER INFORMATION: //function= "approximate span of
? OTHER INFORMATION: acyl carrier domain 2 of module 1"
?
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 6678..11219
? OTHER INFORMATION: //function= "approximate span of
? OTHER INFORMATION: module 2"
?
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 6678..8066
? OTHER INFORMATION: //function= "approximate span of
? OTHER INFORMATION: beta-ketoacyl ACP synthase of module 2"
?
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 8262..9305
? OTHER INFORMATION: //function= "approximate span of
? OTHER INFORMATION: acyltransferase domain of module 2"
?
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 9906..10454
? OTHER INFORMATION: //function= "approximate span of
? OTHER INFORMATION: beta-ketoreductase domain of module 2"
?
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 10707..10964
? OTHER INFORMATION: //function= "approximate span of
? OTHER INFORMATION: acyl carrier domain of module 2"
?
US-08-439-009A-1
Query Match 3.3%; Score 29.4; DB 3; Length 11219;
Best Local Similarity 51.1%; Pred.No. 48;
Matches 69; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 570 gaggaagaacaccatcgctccgcccagaccgtcccttcctcaagaaggcgggaag 629
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2129 GCGCAGGTCACATCGCCCTGCTTACGACCACCCGACTCCGCTGGCGGTGCAGAAC 2188
QY 630 cgltaggggtgatcttcttcgcgaacctgcctatacatgaagcctgcygcaaatggc 689
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2189 ACTCGCGGGGGAACCGAGGTGCACAAAGGAACCGCGCGGACCAAGAGCGC 2248
QY 690 acagctcaaggcgt 704
| | | | | | | | | |
Db 2249 GCCGCGGACCGGT 2263
RESULT 39
US-08-905-223-196
Sequence 196, Application US/08905223
Patent No. 6222029
GENERAL INFORMATION:
APPLICANT: Edwards, Jean-Baptiste D.
APPLICANT: Duclert, Aymeric
APPLICANT: Lacroix, Bruno
TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
NUMBER OF SEQUENCES: 503
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:

```


STATE: Oregon
COUNTRY: United States of America
ZIP: 97204
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3-1/2 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: Wordperfect 5.1/PC Gene
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/914,282D
FILING DATE: July 13, 1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/856,015
FILING DATE: March 23, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Ramon A. Kiltzke II
REGISTRATION NUMBER: 30,188
REFERENCE/DOCKET NUMBER: 2815-36746
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391
TELEFAX: (503) 226-9446
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1432 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double-stranded
TOPOLOGY: Linear
MOLECULE TYPE: Genomic DNA
DESCRIPTION:
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: Flavobacterium sp. Strain ATCC 39723
US-07-914-282D-3

Query Match 3.3%; Score 29.2; DB 1; Length 1432;

Best Local Similarity 49.4%; Pred. No. 18;
Matches 76; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 306 caaagcggtacacgagctgtggaacatgtgcagcagcgtttggacaacacgaagcgtgtct 365
DB 879 CACGGTGGCCAGCGGAGTGCACCTGCAGAAAGCGCTCGAGCACTGACCTGCATCT 938
QY 366 atcatcagcgcacatcgacatgacatgttggcggaagcgtacatcagccagcagct 425
DB 939 GCGGACAAAGCCGTTTCATCGCGGGCAGCAACTACACCATCGCCGACATCATGTGACCGT 998
QY 426 tcggtccgcgtgacgcacatgtacaacacgcg 459
DB 999 CCGTGGCGCGGATGAGATGCTCAACATGACG 1032

RESULT 41
US-08-276-887A-3
Sequence 3, Application US/08276887A

GENERAL INFORMATION:
APPLICANT: Orser, Cindy S. and Xun, Luying
TITLE OF INVENTION: GENES AND ENZYMES INVOLVED
TITLE OF INVENTION: IN THE MICROBIAL
TITLE OF INVENTION: DEGRADATION OF
TITLE OF INVENTION: PENTACHLOROPHENOL
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ramon A. Kiltzke II
STREET: One World Trade Center
STREET: 121 S.W. Salmon Street
STREET: Suite 1600
CITY: Portland
STATE: Oregon

COUNTRY: United States of America
ZIP: 97204
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3-1/2 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: Wordperfect 5.1/PC Gene
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,887A
FILING DATE: July 18, 1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/914,282
FILING DATE: July 13, 1992
APPLICATION NUMBER: 07/856,015
FILING DATE: March 23, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Ramon A. Kiltzke II
REGISTRATION NUMBER: 30,188
REFERENCE/DOCKET NUMBER: 2815-36746
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391
TELEFAX: (503) 226-9446
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1432 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double-stranded
TOPOLOGY: Linear
MOLECULE TYPE: Genomic DNA
DESCRIPTION:
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: Flavobacterium sp. Strain
ATCC 39723
US-08-276-887A-3

Query Match 3.3%; Score 29.2; DB 1; Length 1432;

Best Local Similarity 49.4%; Pred. No. 18;
Matches 76; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 306 caaagcggtacacgagctgtggaacatgtgcagcagcgtttggacaacacgaagcgtgtct 365
DB 879 CACGGTGGCCAGCGGAGTGCACCTGCAGAAAGCGCTCGAGCACTGACCTGCATCT 938
QY 366 atcatcagcgcacatcgacatgacatgttggcggaagcgtacatcagccagcagct 425
DB 939 GCGGACAAAGCCGTTTCATCGCGGGCAGCAACTACACCATCGCCGACATCATGTGACCGT 998
QY 426 tcggtccgcgtgacgcacatgtacaacacgcg 459
DB 999 CCGTGGCGCGGATGAGATGCTCAACATGACG 1032

RESULT 42
PCR-US93-02460-3
Sequence 3, Application PC/TUS9302460

GENERAL INFORMATION:
APPLICANT: Orser, Cindy S. and Xun, Luying
TITLE OF INVENTION: GENES AND ENZYMES INVOLVED IN THE
TITLE OF INVENTION: MICROBIAL DEGRADATION OF PENTACHLOROPHENOL
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: William D. Noonan, M.D.
STREET: One World Trade Center
STREET: 121 S.W. Salmon Street
STREET: Suite 1600
CITY: Portland
STATE: Oregon
COUNTRY: United States of America

```
21P: 97204
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5-inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: Wordperfect 5.1/PC Gene
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/02460
FILING DATE: 19930319
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/856,015
FILING DATE: March 23, 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/914,282
FILING DATE: July 13, 1992
ATTORNEY/AGENT INFORMATION:
NAME: William D. Noonan, M.D.
REGISTRATION NUMBER: 30878
REFERENCE/DOCKET NUMBER: 2815-36746-WDN
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391
TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1432 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: Double-stranded
TOPOLOGY: Linear
MOLECULE TYPE: Genomic DNA
DESCRIPTION:
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: Flavobacterium sp. Strain ATCC 39723
PCT-US93-02460-3

Query Match
Best Local Similarity 3.3%; Score 29.2; DB 5; Length 1432;
Matches 76; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 306 caaagcggtacacggtggaacatgtgcagcagcttgagcaaacagaaagggctct 365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 879 CACGGTGGCCGACGGGAGATGCACTGCAAAAGCGCCTCGACGACTGACGTGCTCT 938

QY 366 attcatcagcgacatggcagctacgattggggcgagcgtctacatcagccagcct 425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 939 GCGGACACAGCCGTTTCATCGCGGCGACCACTACAGCATCGCGCATCATGTGGACCGT 998

QY 426 tccgttcgcgctgacgcgcatgtacaaacgcgcg 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 999 CCTGCTGGCGCGATCGAGATGCTCAACATGAGCG 1032

RESULT 43
US-07-945-283-3
Sequence 3, Application US/07945283
Patent No. 5352596
GENERAL INFORMATION:
APPLICANT: Cheung, Andrew K.
APPLICANT: Wesley, Ronald D.
TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
TITLE OF INVENTION: Involving The Ep0 and L1T Genes
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis P. Ribando
STREET: 1815 No. 5352596th University Street
CITY: Peoria
STATE: IL
COUNTRY: USA
ZIP: 61604
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/945,283
FILING DATE: 19920911
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Ribando, Curtis P
REGISTRATION NUMBER: 27976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 309-685-4011 ext. 513
TELEFAX: 309-685-4128
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1683 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pseudorabies virus
FEATURE:
NAME/KEY: CDS
LOCATION: 211..1440
OTHER INFORMATION: /product="early protein 0"
US-07-945-283-3

Query Match
Best Local Similarity 3.3%; Score 29.2; DB 1; Length 1683;
Matches 58; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 21 gctgttccccccttggaacgcgcacatgcctgttgaccgcctgtcaaatgct 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1500 GTCCTGCCCCCACTGTGGCTCCGCTCTCTCTGCTCGTCCGCGCCCACTGTTCT 1559

QY 81 ctccctgtcgcgcttctctgtctgcacagctgggaaacggctc 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1560 GTTCCTGTTCTTCCGTGTTCTGTTCCCGGTCTGCTGCTCGGCTC 1605

RESULT 44
US-09-321-981-4/C
Sequence 4, Application US/09321981
Patent No. 6287839
GENERAL INFORMATION:
APPLICANT: Genencor International, Inc.
TITLE OF INVENTION: Cellulase Produced Therefrom and Method of Producing Same
TITLE OF INVENTION: Cellulase Produced Therefrom and Method of Producing Same
FILE REFERENCE: GC540-2
CURRENT APPLICATION NUMBER: US/09/321,981
CURRENT FILING DATE: 1999-05-28
PRIOR APPLICATION NUMBER: PCT/US99/11971
PRIOR FILING DATE: 1999-05-28
PRIOR APPLICATION NUMBER: 09/104,308
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 08/974,042
PRIOR FILING DATE: 1997-11-19
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 1716
TYPE: DNA
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Nearest "neighbor" = Streptomyces
OTHER INFORMATION: thermobaceous
US-09-321-981-4
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2002, 08:42:47 ; Search time 27321.5 Seconds
(Without alignments)
1279.881 Million cell updates/sec

Title: US-09-303-518D-465
Perfect score: 1671
Sequence: 1 ttgggcattccgcgcaaaat.....gaaagattacacacaatga 1671

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : GenBml:.*
1: gb.ba:.*
2: gb.htg:.*
3: gb.in:.*
4: gb.om:.*
5: gb.ov:.*
6: gb.pat:.*
7: gb.ph:.*
8: gb.pl:.*
9: gb.pr:.*
10: gb.ro:.*
11: gb.sts:.*
12: gb.sy:.*
13: gb.un:.*
14: gb.vi:.*
15: em.ba:.*
16: em.fun:.*
17: em.hum:.*
18: em.in:.*
19: em.mu:.*
20: em.om:.*
21: em.or:.*
22: em.ov:.*
23: em.pat:.*
24: em.ph:.*
25: em.pl:.*
26: em.ro:.*
27: em.sts:.*
28: em.un:.*
29: em.vi:.*
30: em.htg.hum:.*
31: em.htg.inv:.*
32: em.htg.other:.*
33: em.htg.inv:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	1671	100.0	1671	6	A96432	A96432 Sequence 46
2	1671	100.0	311321	1	NMA322491	AL162754 Neisseria
3	1248.2	74.7	1827	6	AX044009	AX044009 Sequence
4	1248.2	74.7	11786	1	AE002419	AE002419 Neisseria
5	1248.2	74.7	34980	6	AX044030	AX044030 Sequence
6	1231.2	73.7	1887	6	A96430	A96430 Sequence 46
7	1176.2	70.4	2028	6	AX236421	AX236421 Sequence
8	1176.2	70.4	2028	6	AX239731	AX239731 Sequence
9	1176.2	70.4	4425	6	AX236407	AX236407 Sequence
10	1176.2	70.4	4425	6	AX239715	AX239715 Sequence
11	1175.6	70.4	2019	6	AX236447	AX236447 Sequence
12	1175.6	70.4	2019	6	AX239733	AX239733 Sequence
13	1175.6	70.4	2256	6	AX236451	AX236451 Sequence
14	1175.6	70.4	2256	6	AX239737	AX239737 Sequence
15	1175.6	70.4	2421	6	AX236449	AX236449 Sequence
16	1175.6	70.4	2421	6	AX239735	AX239735 Sequence
17	1174.4	70.3	2304	6	AX236465	AX236465 Sequence
18	1174.4	70.3	2256	6	AX239751	AX239751 Sequence
19	1174.2	70.3	2256	6	AX236459	AX236459 Sequence
20	1174.2	70.3	2256	6	AX239745	AX239745 Sequence
21	1174.2	70.3	2421	6	AX236453	AX236453 Sequence
22	1174.2	70.3	2421	6	AX239739	AX239739 Sequence
23	653.4	39.1	683	6	A96426	A96426 Sequence 45
24	294.4	17.6	11700	1	AE002393	AE002393 Neisseria
25	294.4	17.6	349980	6	AX044029	AX044029 Sequence
26	286	17.1	195767	1	NMA722491	AL162758 Neisseria
27	237.2	14.2	10057	1	AE002420	AE002420 Neisseria
28	121.8	7.3	1497	6	A96138	A96138 Sequence 17
29	121.8	7.3	1744	6	A61829	A61829 Sequence 9
30	121.8	7.3	3287	6	A61821	A61821 Sequence 1
31	121.8	7.3	3294	1	AE142582	AE142582 Neisseria
32	110.2	6.4	340806	6	NMA122491	AL162752 Neisseria
33	107.2	6.4	1449	6	A96134	A96134 Sequence 16
34	107	6.4	1452	6	A96132	A96132 Sequence 16
35	107	6.4	10865	1	AE002559	AE002559 Neisseria
36	107	6.4	172325	6	AX044035	AX044035 Sequence
37	98	5.9	375	6	A96130	A96130 Sequence 16
38	51.8	3.1	125020	9	AE429315	AF429315 Homo sapi
39	45.4	2.7	326301	1	AE002514	AE002514 Neisseria
40	45.4	2.7	349980	6	NMA62491	AL162757 Neisseria
41	45.4	2.7	349980	6	AX044033	AX044033 Sequence
42	42.2	2.5	3619	1	SCVALSFP	Y13070 S. coelicolo
43	42.2	2.5	42816	1	SCC88	AL139298 Streptomy
44	42.2	2.5	70423	2	AC102254	AC102254 Mus muscu
45	42.2	2.5	98980	2	AP004215	AP004215 Oryza sat
46	42.2	2.5	186323	2	AP004348	AP004348 Oryza sat
47	41.8	2.5	224448	2	PMAL4P4	AL035477 Plasmodiu
48	40.8	2.4	2653	1	XXGDH	X16399 Gene for gl
49	40.8	2.4	172269	2	AC106508	AC106508 Rattus no
50	40.8	2.4	188489	2	AC016298	AC016298 Homo sapi
51	40.8	2.4	277050	1	AP000989	AP000989 Sulfolobu
52	40.6	2.4	605	8	AF377524	AF377524 Zea mays
53	40.6	2.4	605	8	AF377525	AF377525 Zea mays
54	40.6	2.4	605	8	AF377527	AF377527 Zea mays
55	40.6	2.4	605	8	AF377536	AF377536 Zea mays
56	40.6	2.4	605	8	AF377542	AF377542 Zea mays
57	40.4	2.4	7218	6	I66494	I66494 Sequence 14
58	40.2	2.4	171847	2	AL671007	AL671007 Mus muscu
59	40	2.4	18282	1	AF141932	AF141932 Rhizobium
60	40	2.4	122050	2	AC091512	AC091512 Leishmani
61	40	2.4	347660	2	AP002994	AP002994 Mesorhizo
62	39.8	2.4	390	6	AR135147	AR135147 Sequence
63	39.8	2.4	942	1	AF169562	AF169562 Mesorhizo
64	39.8	2.4	12026	1	AE005109	AE005109 Halobacte
65	39.6	2.4	3386	8	AB012641	AB012641 Mycobacte
66	39.6	2.4	15348	1	AE007163	AE007163 Mycobacte
67	39.6	2.4	47852	1	MTY023	AL022022 Mycobacte
68	39.6	2.4	110000	2	LMFLCHR18_00	AL499630 Leishmani
69	39.6	2.4	146519	2	AC074282	AC074282 Oryza sat
70	39.4	2.4	465	3	AF069170	AF069170 Dyscritul
71	39.4	2.4	162042	2	AC108206	AC108206 Homo sapi
72	39.2	2.3	968	1	ECOPROC	J01665 E.coli proc
73	39.2	2.3	10793	1	AE005758	AE005758 Caulobact

74	39.2	2.3	15311	1	AE007164	Myobacte
75	39.2	2.3	156573	2	AC073075	Homo sapi
76	39.2	2.3	165339	3	CEY57A10A	Caenorhab
77	39	2.3	605	8	AF377520	Zea mays
78	39	2.3	605	8	AF377521	Zea mays
79	39	2.3	605	8	AF377522	Zea mays
80	39	2.3	605	8	AF377523	Zea mays
81	39	2.3	605	8	AF377526	Zea mays
82	39	2.3	605	8	AF377529	Zea mays
83	39	2.3	605	8	AF377530	Zea mays
84	39	2.3	605	8	AF377531	Zea mays
85	39	2.3	605	8	AF377532	Zea mays
86	39	2.3	605	8	AF377533	Zea mays
87	39	2.3	605	8	AF377537	Zea mays
88	39	2.3	605	8	AF377538	Zea mays
89	39	2.3	605	8	AF377540	Zea mays
90	39	2.3	1696	8	ZMU14599	U14599 Zea mays Bz
91	39	2.3	5384	6	AX345110	Sequence
92	39	2.3	36224	1	SCD78	ALU34355 Streptomy
93	39	2.3	163786	2	AC097105	AC097105 Homo sapi
94	38.8	2.3	10732	6	E32986	E32986 Gene encodi
95	38.8	2.3	315000	6	RME603644	AL603644 Rhizobium
96	38.6	2.3	882	6	ARI69213	ARI69213 Sequence
97	38.6	2.3	882	6	BD006386	BD006386 Compounds
98	38.6	2.3	882	6	BD006506	BD006506 Compounds
99	38.6	2.3	1896	1	AF009829	AF009829 Mycobacte
100	38.6	2.3	14707	1	AE006955	AE006955 Mycobacte

ALIGNMENTS

RESULT 1	A96432	1671 bp	DNA	Linear	PAT 07-SEP-2000
LOCUS	A96432	Sequence 465 from Patent WO9924578.			
DEFINITION	A96432				
ACCESSION	A96432				
VERSION	A96432.1	GI:6780108			
KEYWORDS					

SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1671)
AUTHORS Pizza, M., Scariato, V., Rappuoli, R., Grandi, G. and Maignani, V.
TITLE Neisserial antigens
JOURNAL Patent: WO 9924578-A 465 20-MAY-1999;
PIZZA MARIAGRAZIA (IT); SCARLATO VINCENTO (IT); RAPPUOLI RINO (IT);
CHIRON SPA (IT); GRANDI GUIDO (IT); MASNIGANI VEGA (IT)

FEATURES
source 1..1671
Location/Qualifiers
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 484 a 440 c 402 g 345 t
ORIGIN

Query Match 100.0%; Score 1671; DB 6; Length 1671;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ttgggcatccgcgaataatccctattctgtccatctactgtagtgatgctgcgcatg	60
DB	1	TTGGGCAATTCCTCCGCAAAATATCTCTTATCTCTCCATCTGCGAGTGTGCTCCCATG	60
QY	61	catgcacagcctcagatttggcaaacgattctttatccgcgaggttctcgaacgtcag	120
DB	61	CATGCACAGCCTCAGATTGTCGAACGATTTTATCTTCGCGAGGTTCCTCGACCGTCAG	120
QY	121	catcgcacccgcagcaggaataatccctcatctcgcaagcagggggaacttgcagcgc	180
DB	121	CATTCGCAACCCGAGCGGGAATATCACCTATTGGCAGCAGGGGGGAATTTGGCGACGCGC	180
QY	181	agcggtcatatcgatttgggaataacatacaagccatcagttggtgcaaccgttcacacg	240

DB	181	AGCGGTATATTCGATTGGGAAACATACAAAGCCATCATGTTGGGCACTGTTATCCAG	240
QY	241	caggcgccatctaaagaaalatcgctacatlytccgctttccgatacagggcagaa	300
DB	241	CAGCGGCCATTTAAAGGAAATATCGCAATGTGCGCTTTCCGATCACGGGACGAA	300
QY	301	gtccattcccttcgcgaacacatgctccatccgattctgttgtaagccgtagtacc	360
DB	301	GTCCATTCCCTTCGCAACCATGCTTCACATTCGCAATTCGTAAGACCGGTATGCC	360
QY	361	gttgcagatcagccttaccgcatccatttggagagatagcaaacaccatccgcgac	420
DB	361	GTTGACGAGATTACGCTTTACCGCATCATTTGGAGGATTCGAACACCATCCGCCGAC	420
QY	421	ggtatagacggccacagggcgcggtalccgcgtcccaaggcgagagatatatac	480
DB	421	GGTATAGACGGCCACAGGGCGCGGCTATCCGCTCCCAAGCGGAGGATATATAC	480
QY	481	agctacgacataaaggcgctgcccgaataatccgctcaacctgaccggaacacgagc	540
DB	481	AGCTACGACATAAAGCGCTTGCCCAAAATATCCGCTCAACCTGACCGACACCGCAGC	540
QY	541	accggaacagcgtcttgcacggttccacaataccggtatgcttgcagcgaagagta	600
DB	541	ACCGGAACAGCGCTTGTCGACCGCTTCCACATACCGGTATGCTGACGCAAGAGTA	600
QY	601	ggcgacgatttcaaacgcgcacccgatacagcccgagcttgacagatcgggaatgcc	660
DB	601	GGCGACGATTCAAAACGCGCACCCGATACAGCCCGAGCTGACATGCGGCAATGCC	660
QY	661	gccgaagcttcaacggcactgcatatcgtcaaaaacatccgcgcgagcagagaa	720
DB	661	GCGGAGCTTTCAACGCGCACTGCAATATGCTCAAAAACATCATCGCGCGGAGGAA	720
QY	721	atcttgcgcgagcgatgcgtgcgtgcaggtatagaagcgaagctcaaacatctgtatg	780
DB	721	ATTGTCGGCGAGGCGATGCCGTGAGGATTAAGCAAGCAAGCTCAACATGTGCTTATG	780
QY	781	caagcttgggtctgtcttccaccggaacaaagatggtcgagatcaagatttggcagat	840
DB	781	CAGGCTTGCGTCTGTTCCACCGAATAACAAATGCGCGCATCAAGATTGGCGAT	840
QY	841	atgpcgcaactcaaaagactatgctcgacgacgcatccgcatgtggggaagccaaaccc	900
DB	841	ATGGCGCAACTCAAAAGCATATGGCGGACGACCATCCGCAATGGGAGTCCAAAACCC	900
QY	901	aatgcgcacaagcatagaaagcgtcgaacaatatcttaagcagatcaltcccgatcaa	960
DB	901	AATGCCGCAAGCGCATAGAAAGCGCTCAGCAATATCTTTACGCGCATGCCCGTCAAA	960
QY	961	gggatttgaagcgttccggaagaaatacgcttgggggagatcaagcagacatcctgtcaag	1020
DB	961	GGGATTGAGCGTCTTGCGGAAATACGGCTGGGGGATCATCGGCACATCTCTGCACG	1020
QY	1021	cgttcgcagatgtagcagatcgcaltgcggaagaggaatccgcgttcagcgaatatt	1080
DB	1021	CGTTCGCAGATGAGGAGATGCGCATTTGCCGAAGGAAATCCGCGCTCAGCGCAATTTT	1080
QY	1081	ggcgatgcgcatcagcgaataatccgctcccttacatctccggaatatccgttcaaac	1140
DB	1081	GCGGATGCGCATACGCGCAATAATCCCGTCCCTTACCATCTCCGAAATATCCGTTCAAC	1140
QY	1141	ttggagcagcgttaccggaagaaacatcacctcctcaacgcttgcgcgcgttcaacgga	1200
DB	1141	TTGGAGCAGCGTTACGCGCAAGAAACATCACCTCTTCAACGCTGCGCGTCAAAAGGA	1200
QY	1201	aagaatgtgaacttgcgaacaaacgcgacccgaagacccaagtgtccgttttgcgttaaa	1260
DB	1201	AAGAAATGTGAACCTTGCAAAACGCAACCGCAAGACCAAAAGTGTGCTTGAAGGTAA	1260
QY	1261	gggtttccgaatttggaaaagacgttaaatagatagagaattaatccgctgtacaa	1320

Db	1261	GCGTTTCCGAATTTTGAAAAAGACGTAATAATACGATACGAGAAATTAATACCCTGTACA	1320
Oy	1321	caagtgatacctatagatgaaccgcgttcctaactaaaggctctgtagtccgtcat	1380
Db	1321	CAAGTGAAATCCTTAGATGGAACCCGGTCCTTTAATCTTAAAGGTTCTGTGGATCGGCAT	1380
Oy	1381	tcttgatcctaactgcccagaattcaataagcaaatctaccaaaggccaagttagaacga	1440
Db	1381	TCTTGACTTAATCTGACAGATTCAATACGCAAAATTACCAGGCAAGTAGAATCAGA	1440
Oy	1441	tataccaccactaaaataactactcctctccagcacgcgtaccaaaaggaccthaatatga	1500
Db	1441	TATATCCCACTAAAATAATTAATCTCTCTTCAGCACCGCTACCAAAGAACCTAATAATGA	1500
Oy	1501	tatttgataaaatttgttaatgaatgtagcacaaggctccatcaagaataagfcaagaa	1560
Db	1501	TATTTGGATTAATTTGGTTAAAGAAATGAGCATAAAGGTCATCAACACTAAAGGCCAAGAA	1560
Oy	1561	tttgaatggatgtgccatgtctctcaaaacagagaagagacacttggatggctagtagn	1620
Db	1561	TTTGAAATGGAATGTTCAATGTCATAAAGCAGGAAGAGAGCAACTTGGATGGCTAGTAGG	1620
Oy	1621	gatggtaaagcattcaatatatcatcgatgaggaagaagattacacacaaatga	1671
Db	1621	GATGGTAAGCATTTTAAATATATCAATTGATGAGAAAGATTACACACAATAATGA	1671
RESULT 2			
MNA372491			
LOCUS		311321 bp	DNA linear BCT 04-DEC-2000
DEFINITION	NNA372491		
DESCRIPTION	Neisseria meningitidis serogroup A strain Z2491 complete genome;		
ACCESSION	AL162754	AL157959	
VERSION	ALI62754.2	GI:7379424	
KEYWORDS			
SOURCE ORGANISM	Neisseria meningitidis Z2491.		
	Neisseria meningitidis Z2491.		
	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;		
	Neisseria.		
REFERENCE	1 (bases 1 to 311321)		
AUTHORS	Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,		
	Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,		
	Davies,R.M., Davis,P., Devlin,K., Feltham,T., Hamlin,N.,		
	Holtroyd,S., Jagers,K., Leather,S., Moule,S., Mungall,K.,		
	Quail,M.A., Rajandream,M.A., Rutherford,K.W., Simmonds,M.,		
	Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.		
	Complete DNA sequence of a serogroup A strain of Neisseria		
	meningitidis Z2491		
	Nature 404 (6777), 502-506 (2000)		
JOURNAL MEDLINE	20222556		
REFERENCE	2 (bases 1 to 311321)		
AUTHORS	Parkhill,J.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-MAR-2000) Submitted on behalf of the Neisseria		
	sequencing team, Sanger Centre, Wellcome Trust Genome Campus,		
	Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk		
COMMENT	NOTES:		
	Detailed de N. meningitidis sequencing at the Sanger Centre are		
	available on the World Wide Web.		
	(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).		
FEATURES	location/Qualifiers		
source	1..311321		
	/organism="Neisseria meningitidis Z2491"		
	/strain="Z2491"		
	/db_xref="taxon:122567"		
	/note="serogroup: A"		
RBS	gene		
	129..587		
	/gene="NMA0698"		
CDS	129..587		
	/gene="NMA0698"		
	/note="NMA0698, possible periplasmic protein, len: 152 aa		
	unknown, contains a probable N-terminal signal sequence		

and lies within a region of unusually low GC content"

```
/transl_table=1
/codon_start=1
/product="putative periplasmic protein"
/protein_id="CAB83985.1"
/db_xref="GI:7379425"
/db_xref="SPTRMBL:09J002"
/translation="MTKIKILFLFLFLFVLAINILFFPSSDIESFGNYOEYVD
KGWPAIVILKHNIDGNEGNFDRKILSGLVLEYKEDDNIYFSDIGQFASDSCYKKEI
LYGIIILNKKNIININSEKNNFLSEDKIMKGTFRNMLADPKNCKNIQTLD"
742..1077
/genre="NMA0699"
742..1077
/genre="NMA0699"
/note="NMA0699", pseudogene, probable ABC transporter
protein, len: 336 bp, similar to C-termini of many ABC
transporters e.g. TR:Q46973 (EMBL:U47048), mltB,
Escherichia coli microcin transport protein (707 aa),
fasta scores: E(): 2.6e-19, 53.2% identity in 111 aa
overlap. Contains PS00211 ABC transporters family
signature. NMA0686 may be the remainder of this
pseudogene"
/codon_start=1
/pseudo
/transl_table=1
/product="putative ABC transporter protein (pseudogene)"
868..912
/genre="NMA0699"
/note="PS00211 ABC transporters family signature"
complement(1101..2327)
/genre="NMA0700"
complement(1101..2327)
/genre="NMA0700"
complement(1101..2327)
/note="NMA0700", possible ribonuclease BN, len: 408 aa;
similar to SW:RBN_ECOLI (EMBL:L19201), rbn, Escherichia
coli ribonuclease BN (EC 3.1.-.-), fasta scores: E():
2.3e-28, 34.3% identity to TR:085449
at the C-terminus. Also similar to TR:085449
(EMBL:AF067083) Vitreoscilla sp. hypothetical protein (376
aa), fasta scores: E(): 0, 45.0% identity in 333 aa
overlap. Contains a region similar to NMA0677, fasta
scores: E(): 4.3e-08, 64.4% identity in 45 aa overlap.
Contains hydrophobic, possible membrane-spanning regions,
as does rbn"
/codon_start=1
/transl_table=1
/product="putative ribonuclease BN"
/protein_id="GI:7379426"
/db_xref="GI:7379426"
/db_xref="SPTRMBL:09JVT9"
/translation="MTFLIORLGLADNKTICAFMVPVRPDEERYPOAAASMTFTTLL
ALPVLITVAVAASIFPVPRDRSDSFYSFVNOTYPCGADWQVITFNAFRQARLTRA
IGSVMLIVVSLMLIRITDITENRIRKRVNSQRMHQLYVALITFGLSLGICSTSPK
VGSVDPAALASGAPOMSGALRKAATLTFTTLLWGLTVRPKRPVPAKQAVGLATRR
FCELRGRSLFTYMGNFEDGRSIVYGAFVAPFTLLMLNLMTLVGVGVNLSYSYMW
GFEGRFGRSFGRFDFDLKILLLDAAQEGCALPVEFRRIINMGDELGELEKLEKTA
RHGYIYSGGQWVLTKGADSIENLEFKLFYVRPLVRDRHQAQVDAVMMPCTQITLNL
MTAEERDAQKKQOOS"
1376..1385
/genre="NMA0699"
1376..1385
/genre="NMA0699", pseudogene, probable ABC transporter
protein, len: 336 bp, similar to C-termini of many ABC
transporters e.g. TR:Q46973 (EMBL:U47048), mltB,
Escherichia coli microcin transport protein (707 aa),
fasta scores: E(): 2.6e-19, 53.2% identity in 111 aa
overlap. Contains PS00211 ABC transporters family
signature. NMA0686 may be the remainder of this
pseudogene"
/codon_start=1
/pseudo
/transl_table=1
/product="putative ABC transporter protein (pseudogene)"
868..912
/genre="NMA0699"
/note="PS00211 ABC transporters family signature"
complement(1101..2327)
/genre="NMA0700"
complement(1101..2327)
/genre="NMA0700"
complement(1101..2327)
/note="NMA0700", possible ribonuclease BN, len: 408 aa;
similar to SW:RBN_ECOLI (EMBL:L19201), rbn, Escherichia
coli ribonuclease BN (EC 3.1.-.-), fasta scores: E():
2.3e-28, 34.3% identity to TR:085449
at the C-terminus. Also similar to TR:085449
(EMBL:AF067083) Vitreoscilla sp. hypothetical protein (376
aa), fasta scores: E(): 0, 45.0% identity in 333 aa
overlap. Contains a region similar to NMA0677, fasta
scores: E(): 4.3e-08, 64.4% identity in 45 aa overlap.
Contains hydrophobic, possible membrane-spanning regions,
as does rbn"
/codon_start=1
/transl_table=1
/product="putative ribonuclease BN"
/protein_id="GI:7379426"
/db_xref="GI:7379426"
/db_xref="SPTRMBL:09JVT9"
/translation="MTFLIORLGLADNKTICAFMVPVRPDEERYPOAAASMTFTTLL
ALPVLITVAVAASIFPVPRDRSDSFYSFVNOTYPCGADWQVITFNAFRQARLTRA
IGSVMLIVVSLMLIRITDITENRIRKRVNSQRMHQLYVALITFGLSLGICSTSPK
VGSVDPAALASGAPOMSGALRKAATLTFTTLLWGLTVRPKRPVPAKQAVGLATRR
FCELRGRSLFTYMGNFEDGRSIVYGAFVAPFTLLMLNLMTLVGVGVNLSYSYMW
GFEGRFGRSFGRFDFDLKILLLDAAQEGCALPVEFRRIINMGDELGELEKLEKTA
RHGYIYSGGQWVLTKGADSIENLEFKLFYVRPLVRDRHQAQVDAVMMPCTQITLNL
MTAEERDAQKKQOOS"
1376..1385
/genre="NMA0699"
1376..1385
/genre="NMA0699", pseudogene, probable ABC transporter
protein, len: 336 bp, similar to C-termini of many ABC
transporters e.g. TR:Q46973 (EMBL:U47048), mltB,
Escherichia coli microcin transport protein (707 aa),
fasta scores: E(): 2.6e-19, 53.2% identity in 111 aa
overlap. Contains PS00211 ABC transporters family
signature. NMA0686 may be the remainder of this
pseudogene"
/codon_start=1
/pseudo
/transl_table=1
/product="putative ABC transporter protein (pseudogene)"
868..912
/genre="NMA0699"
/note="PS00211 ABC transporters family signature"
complement(1101..2327)
/genre="NMA0700"
complement(1101..2327)
/genre="NMA0700"
complement(1101..2327)
/note="NMA0700", possible ribonuclease BN, len: 408 aa;
similar to SW:RBN_ECOLI (EMBL:L19201), rbn, Escherichia
coli ribonuclease BN (EC 3.1.-.-), fasta scores: E():
2.3e-28, 34.3% identity to TR:085449
at the C-terminus. Also similar to TR:085449
(EMBL:AF067083) Vitreoscilla sp. hypothetical protein (376
aa), fasta scores: E(): 0, 45.0% identity in 333 aa
overlap. Contains a region similar to NMA0677, fasta
scores: E(): 4.3e-08, 64.4% identity in 45 aa overlap.
Contains hydrophobic, possible membrane-spanning regions,
as does rbn"
/codon_start=1
/transl_table=1
/product="putative ribonuclease BN"
/protein_id="GI:7379426"
/db_xref="GI:7379426"
/db_xref="SPTRMBL:09JVT9"
/translation="MTFLIORLGLADNKTICAFMVPVRPDEERYPOAAASMTFTTLL
ALPVLITVAVAASIFPVPRDRSDSFYSFVNOTYPCGADWQVITFNAFRQARLTRA
IGSVMLIVVSLMLIRITDITENRIRKRVNSQRMHQLYVALITFGLSLGICSTSPK
VGSVDPAALASGAPOMSGALRKAATLTFTTLLWGLTVRPKRPVPAKQAVGLATRR
FCELRGRSLFTYMGNFEDGRSIVYGAFVAPFTLLMLNLMTLVGVGVNLSYSYMW
GFEGRFGRSFGRFDFDLKILLLDAAQEGCALPVEFRRIINMGDELGELEKLEKTA
RHGYIYSGGQWVLTKGADSIENLEFKLFYVRPLVRDRHQAQVDAVMMPCTQITLNL
MTAEERDAQKKQOOS"
1376..1385
/genre="NMA0699"
1376..1385
/genre="NMA0699", pseudogene, probable ABC transporter
protein, len: 336 bp, similar to C-termini of many ABC
transporters e.g. TR:Q46973 (EMBL:U47048), mltB,
Escherichia coli microcin transport protein (707 aa),
fasta scores: E(): 2.6e-19, 53.2% identity in 111 aa
overlap. Contains PS00211 ABC transporters family
signature. NMA0686 may be the remainder of this
pseudogene"
/codon_start=1
/pseudo
/transl_table=1
/product="putative ABC transporter protein (pseudogene)"
868..912
/genre="NMA0699"
/note="PS00211 ABC transporters family signature"
complement(1101..2327)
/genre="NMA0700"
complement(1101..2327)
/genre="NMA0700"
complement(1101..2327)
/note="NMA0700", possible ribonuclease BN, len: 408 aa;
similar to SW:RBN_ECOLI (EMBL:L19201), rbn, Escherichia
coli ribonuclease BN (EC 3.1.-.-), fasta scores: E():
2.3e-28, 34.3% identity to TR:085449
at the C-terminus. Also similar to TR:085449
(EMBL:AF067083) Vitreoscilla sp. hypothetical protein (376
aa), fasta scores: E(): 0, 45.0% identity in 333 aa
overlap. Contains a region similar to NMA0677, fasta
scores: E(): 4.3e-08, 64.4% identity in 45 aa overlap.
Contains hydrophobic, possible membrane-spanning regions,
as does rbn"
/codon_start=1
/transl_table=1
/product="putative ribonuclease BN"
/protein_id="GI:7379426"
/db_xref="GI:7379426"
/db_xref="SPTRMBL:09JVT9"
/translation="MTFLIORLGLADNKTICAFMVPVRPDEERYPOAAASMTFTTLL
ALPVLITVAVAASIFPVPRDRSDSFYSFVNOTYPCGADWQVITFNAFRQARLTRA
IGSVMLIVVSLMLIRITDITENRIRKRVNSQRMHQLYVALITFGLSLGICSTSPK
VGSVDPAALASGAPOMSGALRKAATLTFTTLLWGLTVRPKRPVPAKQAVGLATRR
FCELRGRSLFTYMGNFEDGRSIVYGAFVAPFTLLMLNLMTLVGVGVNLSYSYMW
GFEGRFGRSFGRFDFDLKILLLDAAQEGCALPVEFRRIINMGDELGELEKLEKTA
RHGYIYSGGQWVLTKGADSIENLEFKLFYVRPLVRDRHQAQVDAVMMPCTQITLNL
MTAEERDAQKKQOOS"
1376..1385
/genre="NMA0699"
1376..1385
/genre="NMA0699", pseudogene, probable ABC transporter
protein, len: 336 bp, similar to C-termini of many ABC
transporters e.g. TR:Q46973 (EMBL:U47048), mltB,
Escherichia coli microcin transport protein (707 aa),
fasta scores: E(): 2.6e-19, 53.2% identity in 111 aa
overlap. Contains PS00211 ABC transporters family
signature. NMA0686 may be the remainder of this
pseudogene"
/codon_start=1
/pseudo
/transl_table=1
/product="putative ABC transporter protein (pseudogene)"
868..912
/genre="NMA0699"
/note="PS00211 ABC transporters family signature"
complement(1101..2327)
/genre="NMA0700"
complement(1101..2327)
/genre="NMA0700"
complement(1101..2327)
/note="NMA0700", possible ribonuclease BN, len: 408 aa;
similar to SW:RBN_ECOLI (EMBL:L19201), rbn, Escherichia
coli ribonuclease BN (EC 3.1.-.-), fasta scores: E():
2.3e-28, 34.3% identity to TR:085449
at the C-terminus. Also similar to TR:085449
(EMBL:AF067083) Vitreoscilla sp. hypothetical protein (376
aa), fasta scores: E(): 0, 45.0% identity in 333 aa
overlap. Contains a region similar to NMA0677, fasta
scores: E(): 4.3e-08, 64.4% identity in 45 aa overlap.
Contains hydrophobic, possible membrane-spanning regions,
as does rbn"
/codon_start=1
/transl_table=1
/product="putative ribonuclease BN"
/protein_id="
```


Db 143501 ATGGCGCAACTCAAGACATATGCCGAGCAGACCCATCCGATGCGGAGTCCAAACCCC 143560
QY 901 aatgcccacaagcagcatagaagccgtcagcaatacttccaggaatcatcccgctcaaa 960
|||||
Db 143561 AATGCCGACAAAGGCATAGAACCCGTACGCAATATCTTTACGGCAGTCATCCCGTCAAA 143620
QY 961 gggatgtgagctgtctcggggaaaaatacggcttggcgagcatcagcgacatcctgtcaag 1020
|||||
Db 143621 GGGATTGAGAGCTGTTCGGGGAATAATACGGCTTGGGGGCATCAGCGCAATCCTGTCAAG 143680
QY 1021 cggtcgcaatggcgagatcgcatctggcgaagaagaaatccgcgtcagcgacaattt 1080
Db 143681 CGGTCCCAATGGGCGAGATCGGATTTGCCAAAGGAAATCCGCCGTGACGACAAATTTT 143740
QY 1081 gccgatgagcagcatcagcgaataaccgctcccttaacatcccggaatcgcgtttcaaac 1140
Db 143741 GCCGATGGCGGCATACGCCAAATACCCGTCCTTACCATTTCCGAAATATCCGTTCAAC 143800
QY 1141 ttggagcaagcgttaccgcaagaanaacatcacctcctcaaccgttgcgcgtlcaaacg 1200
Db 143801 TTGGACACAGCTTACGCAAAAGAAACATACCTCCTCAACCGTGCCTCAACGGA 143860
QY 1201 aagaatgtgaactgtgcaacaacaacgcccagcgaagaaccgaatgtccgtttgacgtaaa 1260
Db 143861 AAGAAATGTAAACTGGCCAAACCAACGCCACCCGAAAGACCAAAAGTCCGTTTGACGGTAAA 143920
QY 1261 gggcttcgcaatttggaaaaagcgtlaaaatagcatcaggaatttaaccgctgttaca 1320
Db 143921 GGGTTTCCGATTTTAAAAAGCGTAAATATAGCATACGAAATTAATACCGCTGTACCA 143980
QY 1321 caagtgaatcctctagatgaagaccgcttcttaacctaagaagttctgtcgtacgtctcat 1380
Db 143981 CAAGTGAATCCATATAGATGAACCCGCTCTTAATCCATAAGTCTGTCGATCGGCTCAT 144040
QY 1381 tcttgctcctaactgcgcgaatttcaatacgaacaatttaccgaagcgaagttagaatcaga 1440
Db 144041 TCTTGCTCTAATACCTCCGAATTCATATACGCAAAATTAACCAAGGATGATACGA 144100
QY 1441 tatatccacctaanaatactctccttcagcaccgcttaccaaaagaacctaatatagta 1500
Db 144101 TATATCCACACCTAAATAATATCTCTCTTCAGCACCGCTACCAAAAGACCTATATATGGA 144160
QY 1501 tatttggataaatttggtaataatgaatgaactaaagttccatcaagaactaaagtcaagaa 1560
Db 144161 TATTGTGATAAATTTGTATGATGACTAAAGGTCTATCAACAGACTAAAGGTCAAGAA 144220
QY 1561 ttgtgaatggatgttcaatgttctaaacagaagaagagaacttggatggctagtagg 1620
Db 144221 TTTGAATGGGATGTTCAATTTGTCTAAACAGGAAGAGCAACTTGGATGGCTAGTAGG 144280
QY 1621 gatgtgaacatttaataatcatcaattgatggaagaattacacacaataga 1671
Db 144281 GATGTGAACATTTAATATATATCATATTGATGGAAGATTACACACAAATGA 144331

RESULT 3
AX044009 1827 bp DNA linear PAT 24-NOV-2000
LOCUS AX044009
DEFINITION Sequence 88 from Patent WO0066791.
ACCESSION AX044009
VERSION AX044009.1 GI:11342894
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
1 (bases 1 to 1827)
Piazza, M., Hickey, E., Peterson, J., Tetteh, H., Venier, J. C.,
Masignani, V., Galeotti, C., Mora, M., Ratti, G., Scarselli, M.,
Scarletta, V., Rappuoli, R., Prazzer, C. M. and Grandi, G.
Neisseria genomic sequences and methods of their use
Patent: WO 0066791-A 88 09-NOV-2000;
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)

FEATURES
source Location/Qualifiers
1..1827
/organism="Neisseria meningitidis"
/db_xref="taxon:487"
BASE COUNT 539 a 446 c 454 g 387 t 1 others
ORIGIN

Query Match 74.7%; Score 1248.2; DB 6; Length 1827;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1 ttgggcatctcccgcaaaatccctatctcttccactactggtcagtggtcgtcagtg 60
Db 1 TTGGGCAATTTCCCGCAAAATATCCCTTATCTGTCCATCTGCGAGTGTGCTGCCATG 60
QY 61 catgcaacagcctcagatlttggcaaacgaltcttltaccggaaglttccgacgtcag 120
Db 61 CATGCACAGCCCTCAGATTTGGCAAAAGATTCCTTTATCCGCGAGGTTTCGACCGTCAG 120
QY 121 catctgaaccccgagcgggaatatacaccctatctcgcaagcggggggaacttgcgagcgc 180
Db 121 CATTTGGAACCCGACGCGGAATATCAACCTATTTGCGCAGCAGGGGGGAATTTGCCAGCGC 180
QY 181 agcgtcatatcagatltgggaaacatacaagaagcatcagttggcaacctgtcatcag 240
Db 181 ACGGGCATATCGATTTGGGAAAAATACAAAGCCATCATGTTGGCAACCTGATGATTCAA 240
QY 241 caggcgccattaaagaaatalcggctcatctgtccgcttcttcgaatcaagggcagaa 300
Db 241 CAGGGCGCCATTAAAGAAATATCGGCTCATTTGCTCGCTTTCCGATCAGCGGCACGAA 300
QY 301 gtccatctcccttcgaacacacalgcctcacatctcgatctcgtatgaagccgtafccc 360
Db 301 GTCCATTTCCCTTGGACAAACCATGCTCTCATTTCCGATTTCTGATGAAGCGGATGATCC 360
QY 361 gttagcagatctcagccttaccatccatcttggagcagatagcaaacacatcccgccgac 420
Db 361 GTTAGAGGATTTAAGCTTTACGCGATCATCTTGGAGCGAGATACGACACATCCCGCGGAC 420
QY 421 ggccttagcagggcagcagggcgggcgctatcccgctcccaaggcgcgagatataac 480
Db 421 GGCTATGACGGGCGACAGGGCGGGCGGCTATCCCGCTCCCAAAGCGCGGAGATATATAC 480
QY 481 agcttcgaataaagagcgttggccaaataatccgcgtcaacctgacccgcaacccgagc 540
Db 481 ACCTACGACATTAAGAGCGCTTGCCCAAAATATCCGCTCAACCTGACCGCAACCGCAGC 540
QY 541 accggaacaacggctgtcgacagcttccacaataaccggtatgtctgacgaagagta 600
Db 541 ACGGGAACAGCGCTTCCGACACGTTTCCACAAATGCCGGTGTATGCTGACGCAAGGAGTA 600
QY 601 ggcgaacgattcaaacgcgcaccccgatatcagcccgagcttggacagatcgggcaatgccc 660
Db 601 GCGACAGGATTCAAAGCGCGCACCGCATACAGCCCGAGCTGACAGATCGGGCAATGCC 660
QY 661 gccgaagcttcaacgagcactcgagatctcgttaaaaaacatacatcgcgcgcgaagaa 720
Db 661 GCCGAAGCTTTCAACGCGCACTGCGATATCTGTTAAAAACATCATTCGGCGGCGGAGGAA 720
QY 721 attgtcgcgcagcagatccgtgcagaggtatagaaggaaggtcgaacatctgtctatg 780
Db 721 ATTGTGCGCGAGGCGCATCCGTCGAGGGCGATAGGAGAGCTCAACATTTGCTGTATG 780
QY 781 cagcgcttgggtctgtcttccacccgaaaaaagaatgtgcgcatacgaatttggcagat 840
Db 781 CACGGCTTGGGTCTGTCTTCCACCGAAAAACAAGATGGCGGCATCAACGATTTTGGCGAT 840
QY 841 atggcgcaactcaagaactatgtcgcgacgacgacatccgcgatttggcagatccaacccc 900
Db 841 ATGGCGCAACTCAAAAGACATATCCCGACGACGACATCCGCGATTGGGCGAGTCCAAACCCC 900
QY 901 aatgcccacaagcagcatagaagccgctcagcaatatcttaaggcagctcatcccgctcaaa 960

Db	901	AAAGCCGACAGGACATAGAGCCGTCAGCAATATTATGACAGACCAATCCCATCAAA	960	gene	/note="serogroup: B" complement(174. .1139) /gene="NMB0635"
Qy	961	ggagatggagctgttcgggggaaataacgcttggcgagcatacggcacatcctgtcaag	1020	CDS	complement(174. .1139) /gene="NMB0635"
Db	961	GGGATTGGAGCTGTTCGGGGGAAAAATACGGCTTGCGGCGCATCACGACATCTCTAAG	1020		/note="similar to SP:P37247 PID:1197006 percent identity: 61.94; identified by sequence similarity; putative"
Qy	1021	cgttcgagatgtggcgagatcgcatctgcggaagaatccgcgcctcagcgacaatttt	1080		/codon_start=1 /transl_table=11
Db	1021	CGGTCCGAGATGGCGCGCATTCGTCGGAAGGAAATCCGCCCTCAGCGACAAATTTT	1080		/product="transposase, IS30 family"
Qy	1081	gccgatacgagcatacgcgaataaccgcgtcccttaccattcccggaataatcctgtcaac	1140		/protein_id="AAE41058.1"
Db	1081	GGCGATCGGAGATACGCGCAATACCCGTCCTTACCATTTCCGGAATAATCCGTTCAAC	1140		/db_xref="GI:7225864"
Qy	1141	ttggagagcgttaccggaagaataacatcactctcctcaacgctgcgcgctcaaacgga	1200		/translation="MSYQLDGERHYHIOYLSNCHTQVTEIAQJLNRRKSTISREIRH RTQCOQYAEAKAORSRIKORRROPYKIDSLQIHIDPLIRKLSPEQYCAVLCCKHH QTLHSTIYRIYRQDSKNGSTLMOHLRSCSKYRKRYGSTMTRGVNPAVGIENRPA IYDQSRIGDMADRTIVRGOKSALITLIERVRYTIICKIDSLKADVPARAAYBALK AHKDRVHTITDNGKEFFQHTKTRAKAKETPRCRPHYSREKSLNENTNGLINQYPRK QTDFRNISDRIRVODELNHRPKTLGTEYSVLFNLFQPLIH"
Db	1141	TTTGAGACAGGCTTACGGCAAGAAAAATCACTCCTCAACCGTCCGCGCTCAACGGC	1200	gene	complement(1303. .1686) /gene="NMB0636"
Qy	1201	aagaatgtgaactggaacaacagccacccggaagagtcgcgtttgacggttaa	1260	CDS	/note="hypothetical protein; identified by Glimmer2; putative"
Db	1201	AAAAATGTCAAACTGGCAGACCAACCCGGAAGACAGGCGGTACCGTTTGACGGTAA	1260		/codon_start=1 /transl_table=11 /product="hypothetical protein"
Qy	1261	gggtttccgaatttggaaaaagacgtlaaatacgaatacgaataataa	1309		/protein_id="AAE41059.1"
Db	1261	GGGTTTCCGAATTTTGAGAACGACGTGAATATGATACGACACTCGATA	1309		/db_xref="GI:7225865"
RESULT 4	AE002419	11796 bp	DNA	linear	BCT 25-MAY-2000
LOCUS	AE002419	11796 bp	DNA	linear	BCT 25-MAY-2000
DEFINITION	Neisseria meningitidis serogroup B strain MC58 section 61 of 206 of the complete genome.				
ACCESSION	AE002419	AE002098		gene	
VERSION	AE002419.1	GI:7225863		CDS	complement(1744. .3120) /gene="NMB0637"
KEYWORDS	Neisseria meningitidis MC58.				
SOURCE	Neisseria meningitidis MC58				
ORGANISM	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.				
REFERENCE	1 (bases 1 to 11796)				
AUTHORS	Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., Deboy,R., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E., Ciftone,H., Clark,E.B., Cotton,M.D., Uitterback,T.R., Khouri,H., Qin,H., Vamathevan,J., Gill,J., Scariato,V., Maignani,V., Pizze,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R., Rappoli,R. and Venter,J.C.				
TITLE	Complete genome sequence of Neisseria meningitidis serogroup B strain MC58				
JOURNAL	Science	287 (5459)	1809-1815	gene	complement(3139. .4008) /gene="NMB0638"
MEDLINE	20175755			CDS	complement(3139. .4008) /gene="NMB0638"
PUBMED	10710307				/note="similar to PID:1628575 percent identity: 74.13; identified by sequence similarity; putative"
REFERENCE	2 (bases 1 to 11796)				/codon_start=1 /transl_table=11
AUTHORS	Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., Deboy,R., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E., Ciftone,H., Clark,E.B., Cotton,M.D., Uitterback,T.R., Khouri,H., Qin,H., Vamathevan,J., Gill,J., Scariato,V., Maignani,V., Pizze,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R., Rappoli,R. and Venter,J.C.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA				
FEATURES	Location/Qualifiers				
source	1. 11796 /organism="Neisseria meningitidis MC58" /strain="MC58" /db_xref="taxon:122586"				
				gene	complement(4036. .4635) /gene="NMB0639"
				CDS	complement(4036. .4635) /gene="NMB0639"

```
/note="conserved hypothetical protein; identified by  
glimmer2; putative"  
/codon_start=1  
/transl_table=11  
/product="conserved hypothetical protein"  
/protein_id="AAFA1062.1"  
/db_xref="GI:7225868"  
/translation="MSEKPERIVASGNACKLEFGNLFKPYISITVLPSQAFGIPECP  
EPPTEYENALAKARAHAKYSGLPALDDSGCTCAALANGAPIHSARYAGSDPKSDTA  
NMLKLAELVGRKADSCCYCYLVFYFKRKHDPRIIAGVYHGQNDPIPLGONGCYD  
PYLLPHGKTAALDTEVKNRSHRAOALAEILRLAL"  
/complement(4628..4861)  
/gene="NMB0640"  
/complement(4628..4861)  
/gene="NMB0640"  
/note="hypothetical protein; identified by glimmer2;  
putative"  
/codon_start=1  
/transl_table=11  
/product="hypothetical protein"  
/protein_id="AAFA1063.1"  
/db_xref="GI:7225869"  
/translation="MNRKITYLLSVALLFTLAFMLIVLLGAVLLTVGSKAFVAVSLFLFA  
FGALFGQIGSLALILRHKSLRAQSAATKENRYV"  
/complement(4969..5502)  
/gene="NMB0641"  
/complement(4969..5502)  
/gene="NMB0641"  
/note="similar to GB:U42023 SP:P44529 PID:1003152  
PID:1222038 PID:1204383 percent identity: 92.61;"  
identified by sequence similarity; putative"  
/codon_start=1  
/transl_table=11  
/product="inorganic pyrophosphatase"  
/protein_id="AAFA1064.1"  
/db_xref="GI:7225870"  
/translation="MADFNOQLTPGDVGGIINVNEIPAGSNRNIEMNRKLAAPOLD  
RVBPALFAKPTNYGFIPTLDESDDELVLVTEQPLATGVLEAKRIVGVMEFVDGE  
VDKIVCVPADRRNGNNAKTLISDLPOOLINGIIEHFHNYKDLKKAAGTAKVSWGAE  
EAKVIVKESTRNRKKA"  
/complement(5603..6061)  
/gene="NMB0642"  
/complement(5603..6061)  
/note="similar to GB:D10165 SP:P24236 PID:581152  
PID:669113 PID:912430 percent identity: 58.45; identified  
by sequence similarity; putative"  
/codon_start=1  
/transl_table=11  
/product="GATP pyrophosphohydrolase"  
/protein_id="AAFA1065.1"  
/db_xref="GI:7225871"  
/translation="WPMKPLKYPVSALVLYSGDGILLITRTHEGTSWQSYTGLE  
PGETVAOTARREWEETGILLADGLODMHDSVLEYIHHMHRPKGVFENNEHLS  
AEIPRDPIALQPEEHVSYGWFDEMAAEKVPSPNRRAILLELGRFLGR"  
6272..6348  
/gene="tRNA-Pro-3"  
/product="tRNA-Pro"  
6272..6348  
/gene="tRNA-Pro-3"  
/gene="tRNA-Pro-3"  
6506..8332  
/gene="NMB0643"  
6506..8332  
/gene="NMB0643"  
/note="MaB-related protein; identified by Oxford;  
putative"  
/codon_start=1  
/transl_table=11  
/product="MaB-related protein"  
/protein_id="AAFA6213.1"  
/db_xref="GI:7413435"  
/translation="MGISRKISLIISIAVCLPMRAHASDLANDSFITROVLQNHPEP  
DGKYHLFGSRLELAERSGHITGKIOSHOLGNLMTIOOAAIKNGITIVRPSDHGEVH
```

```
SPFDNHAHSDSDSEAGSPVDGESLYRIHMDGYEHHPADVDGDPGGGYPARKGARDIY  
SYDIKVAONIIRLNLTDNRSTGQRLAFRPNFAGNSMLQGVGDGRKATRSPELDRSG  
NAEAFENGSTADIYKNIIGAAGEIVGADAQVSEGSINIYVMBGLGLSTENKMARIN  
DLADMOLKDYAAAIIRDMVAONPNAQGLEAVENIEMAIPIKIGAVRGKYGIGGI  
TAHPIKRSOMGATALAPKSAVDNFAADYAKRPSYHSRNTSINSEORYGKENTIS  
STVPPSNGKYNKVLADQRHPTGYPFEDGKPPNPKHYKXPKLDIOELSGGIPKAKP  
VEPAKPRWEVDRLNKLITREOEKYNVOELRNGININSFNSOHAOLERINKLSNADEI  
NEADGKCFEFDSDMKDRKFSRLVKSNGFTNPVVEYVEINQKAYIVRGNRVEPAEY  
LGRILHELFKVFPEPVNPTSMKNPTDVNLNDSGNVKKRPRYSK"  
8335..8799  
/gene="NMB0644"  
  
Query Match 74.7%; Score 1248.2; DB 1; Length 11796;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 1271; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
  
Qy 1 ttgggaattccgcgcaaatatcccttatcttctgtccatccagcgtgtgcctgcagatg 60  
|||||  
Db 6506 TTGGGCATTTCGCCCAAAATATCCCTTATTCGTCAATCGCGAGTGTGCTCCGATG 6565  
|||||  
Qy 61 catgcacgcctcagattgtgcaaacgattcttaccgcgacggtctcagcgtcag 120  
|||||  
Db 6566 CATGCACAGCCTCAGATTGTGGCAACGATTCCTTTATCCGGCAGTTCGACCGTGAG 6625  
|||||  
Qy 121 cattcgaaccgcgcggaataaccacctatctgcgcacgacgagggggaactgtccagcgc 180  
|||||  
Db 6626 CATTCGAACCCGAGGGGAATACCACTATTCGCACAGGGGGGAATCTCCGAGCCG 6685  
|||||  
Qy 181 agcggtcatatcgatttgggaataataaacgcatcaggttgggaacctgttcacag 240  
|||||  
Db 6686 AGCGGCGATATCGGATTTGGAAAATACAAAGCCATCGTTGGCACTGTGATTCAA 6745  
|||||  
Qy 241 caggcgccatlaaaggaataatcgcgtacatctgtcgccttccgatalcagggcagaa 300  
|||||  
Db 6746 CAGCGCGCATTAAGAAATATCGGTACATTTGCCGTTTCCGATACGGGACGAGAA 6805  
|||||  
Qy 301 gtccattcccccctcgacaacatcgtcctcaatctcgaatctgtatgaagcggtagtccc 360  
|||||  
Db 6806 GTCCATTCCCCCTTCGACAAACCATGCTTCACATTCGATTCGATCAACCCGTAATCC 6865  
|||||  
Qy 361 gttaacgattcagcgttaccgcatcattgtggacggaataaacacatcccgccagc 420  
|||||  
Db 6866 GTTACGGAATTTAGCGCTTTACCGCATTCATTTGGACGGATAGAACACCATCCCGCCAG 6925  
|||||  
Qy 421 ggcatagcgcggccacgaaggcggtatccgcgtccccaaggcgcgaggaatatatac 480  
|||||  
Db 6926 GGCATATGACGGGCCACAGGGCGGGCGGTATCCCGCTCCCAAGGCGCGAGGATATATAC 6985  
|||||  
Qy 481 agctacgacataaaagcgttgcacaaaatcgcgcctcaaccttgacgcgaacgcgcagc 540  
|||||  
Db 6986 AGCTACGACATTAAGGGCGGTGCCCAAAATATCCGCTCAACCTGACCGAACCACCCACAG 7045  
|||||  
Qy 541 accggaacagcgttgcgcacgttccacaatatcgcggaatgatatgtgcgaagaagta 600  
|||||  
Db 7046 ACCGGAACAGCGCTTGCACGCCGTTTCCCAATAGCCGAGTAGTAGTGACGCAAGAGTA 7105  
|||||  
Qy 601 ggcgcggaatcaaacgcgcgcacccgatacaagcccagcgttgacacgatccgggcaatgccc 660  
|||||  
Db 7106 GCGGACGGATTAACAGGGCCACCGGATACAGCCCGAGCTGGACAGATCGGGCAATGCCC 7165  
|||||  
Qy 661 gccgaagcttcaacgcgcacgtgcagatatcgtcaaaaacatcatcggcgcggaagaaga 720  
|||||  
Db 7166 GCCGAAGCCTTCACAGGCGACTGTCAGATATCGTTAAAMATATATCGGCGCGGACGAGAA 7225  
|||||  
Qy 721 attgtcgcgcgcgcgcgcgtgcaggtatgaagcgaaggtccaacattgctgtatg 780  
|||||  
Db 7226 ATTGTGCGCGCGAGCGGATCCGTGACGGGCATTAAGCGGAAGGCTCAACATTTGCTGCATG 7285  
|||||  
Qy 781 caagcgttgggtctgtcttccacgcgaaaaacgaatgcgcgacatcaagatttggcagat 840  
|||||  
Db 7286 CACGCGTTGGGTCTGCTCTTCACACGAAACAGATGGCGCGCATCAAGATTGTCGAGAT 7345  
|||||  
Qy 841 atgycgcaactcaagaactatgctgcgcagcagccatcgcgattgtggcagtlccaaaacccc 900
```

```
|||||
Db 7346 ATGGCGCAACTCAAGACTATGCGCAGACGATCCGGATTGGGGAATCCAAACCC 7405
QY 901 aatgccgcaagaagatgaagcgttcagcaatatcttaagcaatcattcccgtaaa 960
Db 7406 AATCCCGACAAAGGATATGAAACCGCTCAGCAATATCTTATGACGACCAATCCCATCAA 7465
QY 961 ggaattgagctgttcgggaaatacagcttggcgagcatcagcgcaatcctgtcaag 1020
Db 7466 GGGATTGGAGCTGTTCGGGAAATACGGCTTGGCGGCATCCGACATCTCTATCAAG 7525
QY 1021 cggtcgcaagatggcgagatcgcatctgcgaaagggaatccgcgtcagcgcaattt 1080
Db 7526 CGGTGCGAGATGGCGCGCATGCGATTGCGGAAGGGAAATCCCGCTCAGCGCAATTTT 7585
QY 1081 gccgatcgagatacgcgaataaccgctcccttaccattcccgaaatattcggtcaac 1140
Db 7586 GCCGATGCGGATACGCCAATATCCCGTCCCTTACCATTTCCGAAATATCCGTTTAAAC 7645
QY 1141 ttgagcagcgttacgcgcaaaacacatcacctcctcaacgctgcgcgtcaaacgga 1200
Db 7646 TTGGAGCAGCGTTTACGCAAGAAACATACCTCTCTCAACGCTGCGCGTCAAAAGGC 7705
QY 1201 aagaattgaaactcggcaaaacacgcccgaagaagccaagtcggttgaagtaaa 1260
Db 7706 AAAAATGTCAAACTGGCAAGCCCAACGCCAAGCAAGCGCTACCGTTTGAACGGTAAA 7765
QY 1261 gggttccgaatttgaanaagactaaatacgaatacgaagaattaa 1309
Db 7766 GGGTTTCCGAATTTTGAACACACCTGAAATTTGATACGAACCTCATTA 7814

RESULT 5
AX044030 349980 bp DNA linear PAT 24-NOV-2000
LOCUS Sequence 109 from Patent WO0066791.
DEFINITION AX044030
ACCESSION AX044030
VERSION AX044030.1 GI:11342914
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis.
Neisseria meningitidis
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE
1 (bases 1 to 349980)
Pizze,M., Hickey,E., Peterson,J., Tettelin,H., Venier,J.C.,
Massignani,V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M.,
Scariato,V., Rappunoli,R., Frazer,C.M. and Grandi,G.
Neisseria genomic sequences and methods of their use
Patent: WO 0066791-A.109 09-NOV-2000;
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
Source
1. 349980
/organism="Neisseria meningitidis"
/db_xref="taxon:487"
/notes="sequence too long, cut in 8 pieces. -seq 1: 1 to
349980 349980 bases-seq 108: 300001 to 649980 349980
bases-seq 109: 600001 to 949980 349980 bases-seq 110:
900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980
349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq
113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to
2272325 172325 bases"
BASE COUNT 82523 a 82940 c 96712 g 87805 t
ORIGIN
```

```
Query Match 74.7%; Score 1248.2; DB 6; Length 349980;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
```

```
QY 1 ttggagcattccgcgaataatccctattctgtcatatcagtcagtcgttgcgtcgatg 60
|||||
Db 75614 TTGGGCAATTTCCCGCAAAATATCCCTATTCTGTGCATACAGGCGAGTGTGCTGCCGATG 75673
```

```
QY 61 catgcacagcctcaagatttggcaaaagattctttatccggcaggttctcgaccgtcaag 120
|||||
Db 75674 CATGCACAGCCTCAATTTGGCAAAAGATTCCTTTATCCGGAGGTTCTCGACCGTAG 75733
QY 121 catttgaaccccgagcggaataaccacttctgcgcaagggggaaacttgcgagcgc 180
|||||
Db 75734 CATTTGAAACCGAGCGGAATATCCACCTATTTCGGACAGCAAGGGGGAATCTTCCGACGGC 75793
QY 181 agcgttcataatcggatttgggaaacatacaaaagccatacgaatttggcaactgttcaag 240
|||||
Db 75794 ACCGGCATATCGGATTTGGGAAATATCAAAAGCATATGTTGGGCAACCTCATGATTTCA 75853
QY 241 caggcgccattaaagaaataatccgctacatgttcgcttccgattccgatacaggcaaa 300
Db 75854 CAGCGCGCCATTAAAGAAATATCGGCTACATTGTCCGCTTTTCCGATACGGGCACGAA 75913
QY 301 gtccattcccccctgcacaacacatgctcctcaatctcgattctgaatgaagccgttagtccc 360
Db 75914 GTCCATTCCCTTCGACAAACATGCTTCACATTCCGATTCTGATGAAAGCGGTATGTC 75973
QY 361 gttgacgattcagcctttacccatccatcttggagagatacgaacaacatccgcgac 420
Db 75974 GTTGACGATTTAGCTTTTACCGCATTCATTGGAGCGGATACGAACACCATCCGCGCAG 76033
QY 421 ggtctagcggccacagggcgcggtctatcccgctcccaaggcgcgagagatatatc 480
Db 76034 GGCATTAGACGGGCGACAGGGCGGGCGCTATCCCGCTCCCAAGCGCGAGGATATATAC 76093
QY 481 agctacgacataaaagcgttgcacaaataatccgcctcaactgcacgacaacccgcagc 540
Db 76094 AGCTACGACATAAAGGCGCTTGGCCAAATATCCGCTCAACCTGACCGACACCGCGACG 76153
QY 541 accggacaacggccttgcagcgtttcccaataacggtatgttgcgcgcaagagatga 600
Db 76154 ACCGGACAACGGCTTGCACACCGTTTCCAAATGCGGTATGCTGACGCAAGGAGTA 76213
QY 601 ggcgacgattcaaaacgcgcacaccgcatacagcccgagcttgcagagatcgggcaatgcc 660
Db 76214 GCGCAGGATTCAAAGCGCCACCCGATACAGCCCGAGCTGGACAGATCGGGCAATGCC 76273
QY 661 gccgaagcttcaacggacgtcagatcagatcgttaaaaaactatccgcgcgagaggaa 720
Db 76274 GCGGAAGCCTTCAACGGCGCCTCAGATATCGTAAAAACATCATGCGCGCGCGAGGAA 76333
QY 721 atttgcgcgcagcagatccgtcgtcaggttatagcgaaggtccaacattcgtgtatg 780
Db 76334 ATTGCGGCGCAGGGGATGCCGTGCGAGGCGATACGGAAGGCTCAAAACATTGCTGTATG 76393
QY 781 caagcgttggtctgcttccacgcgaaacaagaatggcgcatcaacgaatttggcagat 840
Db 76394 CACGGCTTGGGCTCTCTTCCACCGAAACAGATGGCGCGCATCAACGATTTGGCAGAT 76453
QY 841 atgctgcgaactcaaaagactatgccgacgaagcattccggaatttggggaatccaaacccc 900
Db 76454 ATGGCGCACTCAAAAGACTATCCGCAAGCAGCATCCGCAATTGGCGAGTCCAAACCCC 76513
QY 901 aatgcgcgaacagatagaagcgttcaacatatcttlaagcagtcattcccccgttaaa 960
Db 76514 AATGCGCACAAAGCATATGAAGCCGTGACCAATATTTTATGCGACGACATCCCATAAA 76573
QY 961 ggaattgagctgttcgggaaatacagcttggcgagcatcagcgcaatcctgtlcaag 1020
Db 76574 GGGATTGGAGCTGTTCGGGAAATATACGGCTTGGGGGCGCATCAAGGCGCATCTATCAAG 76633
QY 1021 cggtcgcaagatggcgagatcgcatctgcgaaagggaatccgcgttcagcgcaattt 1080
Db 76634 CGGTGCGAGATGGGCGCATGCGATTGCGGAAGGAAATCCGCGCTGACGCAATTTT 76693
QY 1081 gccgatcgagatacgcgaataaccgctcccttaccattcccgaaatattcggtcaac 1140
Db 76694 GCCGATGCGGATACGCCAATATCCCGTCCCTTACCATTTCCGAAATATCCGTTCAAC 76753
QY 1141 ttggaagcgttacgcgcaaaacacatcacctcctcaaccgctgcgcgtlcaaacgga 1200
```

Db	76754	TTGAGCAGCGCTTACGGCAAGAAACATCACTTCCTCAACCGTCCCGCTCAACGGC	76813
QY	1201	aagaatgtgaaactctgcacaacaacgccaccggaagaccgaagtcggtttgacggttaa	1260
Db	76814	AAAATGTCAAACTGCGCAGACCAACGCCACCCGAAGACAGCGGTACGCTTGACGCTTAA	76873
QY	1261	gggttccggaattttgaaaaagacgtaaataacgatacagagaataata	1309
Db	76874	GGGTTCCGAAATTTTGAGACACGCTGAATAATGATCAGAAGCTCGATA	76922
RESULT	6		
LOCUS	A96430	1887 bp	DNA linear PAT 07-SEP-2000
DEFINITION	Sequence 463 from Patent WO924578.		
ACCESSION	A96430		
VERSION	A96430.1	GI:6780107	
KEYWORDS			
SOURCE	unidentified.		
ORGANISM	unidentified.		
REFERENCE	unclassified.		
AUTHORS	1 (bases 1 to 1887)		
TITLE	Pizza, M., Scarlato, V., Rappuoli, R., Grandi, G. and Masiugnani, V.		
JOURNAL	Neisserial antigens		
	Patent: WO 924578-A 463 20-MAY-1999;		
	P1ZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);		
	CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)		
FEATURES	Location/Qualifiers		
source	1. 1887		
	/organism="unidentified"		
	/db_xref="taxon:32644"		
BASE COUNT	554 a 468 c 493 g 371 t		1 others.
ORIGIN			

Query Match	73.7%	Score 1231.2	DB 6	Length 1687
Best Local Similarity	96.3%	Pred. No. 6.le-311		
Matches 1260	Conservative 0	Matches 49	Indels 0	Gaps 0
Oy	1	ttggagcattccggcaaaatcatcccttatctctgtcatactgcagtgagctgcagatg	60	
Db	1	TTGGGCAATTTCCCGCAAAATATCCCTTATCTGTCCATATGCGAATGTCCTGCGGAG	60	
Oy	61	catgacacacgcctcagatttggcaaacagatcttcttatcccgaggttcgcacgtcag	120	
Db	61	CATGACACAGCCTCAGATTTTGGCAACAGATCCCTTATTCGGGCAAGTTCTCGACGCTAG	120	
Oy	121	catttcgaaccgcagcgaggaatacacactatctcgcgcagcaggggggaacttgcagagc	180	
Db	121	CATTTCGAACCCGAGGGGAATATCCACCTATTTCCGCGACGAGGGGGAGCTTCCNAGCGC	180	
Oy	181	agcgatcatatcggatttgggaaacatacaaaccatcagtttggcaactgttcatccg	240	
Db	181	AACGGCCAAATGGAATTTGGGAAACATATCAAAACCAATCACTTTGGCCACTCGATGATTCAA	240	
Oy	241	caggcgcccatlaaaggaaatactgcctacatctgtccggttttcgcatacggcagca	300	
Db	241	CAGGGCGCGCTTGAAGGAATATCGGCTATCTTGCTCCGCTTTCCGATACAGGGCACAA	300	
Oy	301	gtccattcccccctgcgaacacatgacctacatccgattctgataagccggttagtccc	360	
Db	301	TTTCCATTTCCCTTGACAAACCATGCTCTCATTTCCGATTTGACGAAGCGGTAGTCCC	360	
Oy	361	gtttgacgattcagccttaccgcatcatttggagcgcgatacgaacacatccgcgcgac	420	
Db	361	GTTTGACGATTTACGCTTTACCGCATATTCGGACATTTGGGACGGATACCAACCATCCCGCGAC	420	
Oy	421	ggccttgaaggcgcacagcgcgcggtctatcccgctcccaaaagcgcgaggatataac	480	
Db	421	GGCTTATGACGGGCGCACAGGGCGGCGCTATTCGCGTCCCAAAAGGCCGAGGATATATAC	480	
Oy	481	agctacgacataaaagcgtttgcaccaaatatctcgcttcaaccttgcgcgacaacgcgac	540	

Db	481	AGTACGACATAAAAAGGCTTGCCCAAAATATCCGCGCTCAACCTGACGACCAACCGCAGG	540
Oy	541	acggacaacagcgtctgttcgacccgtttccacaataacggttagtattgctcgacgcaaggagta	600
Db	541	ACCGGACAAACGGCTTGCGCGACCGCTTCCACAAATGCCCGCGCTATGCTGACCAAGAGATA	600
Oy	601	ggcgacggaattccaacgcgccaccggaatacagccccggagctcgagacagatcgagcaatgccc	660
Db	601	GGCGAGGATTTCAACGGCCACCGCATACGCCGCCGACGCTGGACAGATCGGGCATGCC	660
Oy	661	gccgaagctttcaacgagcactgcatatactgcaaaaacatcatcgagcgcgagagagaa	720
Db	661	GCCGAAGCCTTCAACGGCATCTGCAGATATCTGTCAAAAAATCATATCGCGCGCGAGAGATA	720
Oy	721	attgtcgagcgagcgagatgcccgttcgaaggtattaaagcgaaagctaaatattgctgltatg	780
Db	721	ATTGTGCGCGAGGCGAGGATCGCTGTGAAGGCTATTAACGGAAGCTCAAAATGCTGTCTATG	780
Oy	781	caacgcttggtgtcttgcctttccaccgaaacaagaatgtagcgcatcaagaatttggcagat	840
Db	781	CACGCGCTTGGGCTGTCTCTTCCACGCAAAAAACAAGATGCCGCATCAAGATTTGGCAGAT	840
Oy	841	atggcgcaactcaaaagactatgctcgagcagccatcgcgatttggcgagtcaccaaacccc	900
Db	841	ATGGCGCAACTCAAAAGACTATATGCCGACACGACCACTCCGATTTGGGAGTCCAAAACCCC	900
Oy	901	aatggcgacaagaagatagaagcgtcaagcaatattcttaacgagcatatcccccgtcaaa	960
Db	901	AATGCCGCACAAGCATATGGAAGCGCTCAGCAATATCTTTATGCGAGCATATCCCATCAA	960
Oy	961	gggaattggagctgtctcgggagaaatacggcttggcgagcatcaacgacatctctgtcaag	1020
Db	961	GGGATTGGAAGCTGTCTGGGGAAAAATACGGCTTGGCGGCGATCACGGCAATCTCTCAAG	1020
Oy	1021	cggttcgacgaattggggagaatcgcatttgcgcgaaggaaatccgcgttaagcgacaattt	1080
Db	1021	CGGTTCGAGATGGCGGCATCGATTGCGGAAGGGAAATCCGCCGTACGGCAGACAAATTTT	1080
Oy	1081	gccgaatgcgacatacgccaataaccgctcccttaccatctcccgaaatactcgttcaaac	1140
Db	1081	GCGGATGGCGGATACGCCAAATACCCGTCCTTACATTCGCCGAATATGCGTGTCAAAC	1140
Oy	1141	ttggaagcagcgttaccgcaaaagaacatcaccttcctaaacgctgcgcgttcaagga	1200
Db	1141	TTGGAGCAGCGCTTACGGCAAAAGAAAACATCACTCTCTCAACCGTCCGCCGCAAAACGGC	1200
Oy	1201	aagaatttgaaaacttggcaacaacagccaccccggaagacaagaatgctgcgttttgaaggttaa	1260
Db	1201	AAAAATGTCAAACTGGCAGACCAACGCCACCCGGAAGACAGGGTACCGTTTACGCTAAA	1260
Oy	1261	gggttcgcgaatttgaaaaaagcgtlaaaatacgaatacgaagaataata 1309	
Db	1261	GGGTTCGCAATTTTGAGAGACACGTTGAAAATATGATACGAACCTCGATA 1309	
RESULT	7		
LOCUS	AX236421	2028 bp	DNA linear PAT 26-SEP-2001
DEFINITION	Sequence 114 from Patent WO0164922.		
ACCESSION	AX236421		
VERSION	AX236421.1	GI:15796033	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	Arico,M.B., Comanducci,M.C., Galeotti,G.C., Masignani,V.C.,		
	Guiliani,M.M. and Piza,M.C.		
	Heterologous expression of nonserial proteins		
	Patent: WO 0164922-A 114 07-SEP-2001;		
TITLE			
JOURNAL	Chiron Spa (IT)		
FEATURES			
	Location/Qualifiers		

```
source 1. .2028
        /organism="synthetic construct"
        /db_xref="taxon:32630"
        /note="deltaG741-ORF46.1"
BASE COUNT 557 a 577 c 531 g 363 t
ORIGIN

Query Match 70.4%; Score 1176.2; DB 6; Length 2028;
Best Local Similarity 97.3%; Pred. No. 1.5e-296;
Matches 1196; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 71 cctcagattggcaaaagatcttltatccggcaggtctcgaacgtcagcatttcgaac 130
   |||
Db 773 CCTCAGATTGGCAAAAGATCTTATCCGCGACGTTCTCCACCGCTCAGCATTTTGAAC 832

QY 131 ccgacggaaataccactattcggcagcagggggaacttcggcagcagcaggtcaca 190
   |||
Db 833 CCACGCGAAATACCCCTATTCCGACAGAGGGGGAACCTTCCGACGCGCAAGCGCCATA 892

QY 191 tcggattgggaaacatacaaaagccatcagttgggcaacctgttcacacgagcgagcca 250
   |||
Db 893 TCGGATTGGGAAAAATACAAGCCATAGTTGGGCAACCTGATGATTCACAGCGGCCCA 952

QY 251 ttaaggaataatcgtgctacattgtccgcttccgatacgggacgaagtcacatccc 310
   |||
Db 953 TTAAGGAATAATCGTGCTACATTGTCCGCTTTCCGATCAGCGGACGAAGTCATTCCC 1012

QY 311 ccttcgaaacacatcgtctacatccgattctgtatgaagccggtgtcccgattgaagcat 370
   |||
Db 1013 CCTTCGAACACATGCTCTCATTCGATTCTGATGAAAGCCGCTGATCCGTTGACGGAT 1072

QY 371 tcagccttaccgcacatcatttggagcagatatacgaacacatcccgccgacgctatgacg 430
   |||
Db 1073 TTAGCCTTTACCGCATTCATTGGGACGATAGCAACACATCCCGCGACGCTATGACG 1132

QY 431 ggcacacagggcgcgctatcccgctcccaaaagcgcgagggatatacagctacgaca 490
   |||
Db 1133 GGCCACAGGGCGGGGCTATCCCGCTCCCAAAAGCGCGCAGGATATATACACTAGACAA 1192

QY 491 taaagcgctgcccacaaataatccgctcaacactgacgcgaacacgcgcagcagcaac 550
   |||
Db 1193 TAAAGCGCTGGCCCAAAATATCCGCTCAACCTGACGACGACACCGCACCGGACAAC 1252

QY 551 ggccttgcagccgtttccacaataccggtatgtatgtcgaacgaagatlagcgacgcat 610
   |||
Db 1253 GCCTTGCAGCAGCGTTTCCACAATGCGGTAGTATGCTGACGGAAGAGTGAAGCGAT 1312

QY 611 taaagcgccaccacgatacagcccgagctggaacatcgggcaatgcgcggaagctt 670
   |||
Db 1313 TCAAAAGCGCGCACCCCATATACGCCCGAGCTGAGACAGATCGGGCAATGCGCGCAAGCCT 1372

QY 671 tcaacggcactgcagatatcgttcaaaacatcatcgcgcgcgcaagagaatgttcgcgcg 730
   |||
Db 1373 TCAACGGCACTGCAGATATGTTTAAACAATCATTCGCGCGGACGAGAAATTTGTCGGCG 1432

QY 731 cagcgagatgcgtgcaggggtataagcgaaaggtcctaacaattgtctatgcagcgcttgg 790
   |||
Db 1433 CAGCGATGCGCGTGCAGGCGATTAAGCGAAGCGCTCAAAACATTGCTCATGACGCGTTGG 1492

QY 791 gtcgtgttccacgaaataaagaatgctgcgcatcacaagatttggcagatatgtgcgcaac 850
   |||
Db 1493 GTCTGCTTTTCCACCGAAAAAATAATGGCGGCATCAACGATTTGGCAATATATGGCGCAAC 1552

QY 851 tcaaaactatgcgcgacgagcagcatcgcgcatitggcaggtccaaccacatgcccagc 910
   |||
Db 1553 TCAAAACATATGCGCGACGACGACCATCGCGATTTGGCGATCCAAAACCCCAATGCGCGCAC 1612

QY 911 aagcgatagaagcgctcagcaataatcttaccgcaagtcacatcccgctcaaaaggaattggag 970
   |||
Db 1613 AAGCGATAGAAGCGCTCAGCAATATCTTTATGCAAGCAGCATCCCATCAAAAGGATTTGGAG 1672

QY 971 ctgttcggggaaataatcgcgtcttggcgcgcatcacaagcacaatccctgtcaagcggtcgcaaga 1030
```

```
|||||
Db 1673 CTGTTCCGGGAAATATACGGCTTGGGGGCAATCAGGCACATTCATACAGGGGTCCGACA 1732

QY 1031 tggcgagatcgcatttgcgcgaagaaggaatccgctcagcgacacatttgcgcatgcg 1090
   |||
Db 1733 TGGCGGCGATGCCATTGGCCGAAAGGGAATCCGCCGTCACGACGACAAATTTTCCCATGGG 1792

QY 1091 catagccaataaccgctccctaccatcccggaatatccgttcaacttggagcagc 1150
   |||
Db 1793 CATAGCCCAATATCCCGTCCCTTACCATTCCCGAAATATCCGTTCAAACTTGGAGCAGC 1852

QY 1151 gttacgcaaaagaaacatacactcctcaacgcgttgcgcgcgttcaaacggaagaatgtga 1210
   |||
Db 1853 GTTACGCAAAAGAAAACATCACCTCTCAACCGTGC CGCGCTCAACAGCGCAAAATGTCA 1912

QY 1211 aactggcaacaagccaccccggaagacccaagatgcgcttgcagcgttaaggggttccga 1270
   |||
Db 1913 AACTGGCAGACCAACGCCACCCGAAAGACAGGCGGTACCGTTTGAGGTTAAAGGTTTCCGA 1972

QY 1271 atttgaaaaagacgtaaaatacgaacg 1299
   |||
Db 1973 ATTTGAGACAGCACGTGAATATGATACG 2001

RESULT 8
AX239731 2028 bp DNA linear PAT 26-SEP-2001
LOCUS AX239731
DEFINITION Sequence 33 from Patent WO0164920.
ACCESSION AX239731
VERSION AX239731.1 GI:15797376
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE
1 (bases 1 to 2028)
AUTHORS
Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
and Piza,M.
TITLE
Hybrid expression of neisserial proteins
JOURNAL
Patent: WO 0164920-A 33 07-SEP-2001;
Chiron Spa (IT)
FEATURES
Location/Qualifiers
source 1. .2028
        /organism="synthetic construct"
        /db_xref="taxon:32630"
        /note="deltaG741-ORF46.1"
BASE COUNT 557 a 577 c 531 g 363 t
ORIGIN

Query Match 70.4%; Score 1176.2; DB 6; Length 2028;
Best Local Similarity 97.3%; Pred. No. 1.5e-296;
Matches 1196; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 71 cctcagattggcaaaagatcttltatccggcaggtctcgaacgtcagcatttcgaac 130
   |||
Db 773 CCTCAGATTGGCAAAAGATCTTATCCGCGACGTTCTCCACCGCTCAGCATTTTGAAC 832

QY 131 ccgacggaaataccactattcggcagcagggggaacttcggcagcagcaggtcaca 190
   |||
Db 833 CCACGCGAAATACCCCTATTCCGACAGAGGGGGAACCTTCCGACGCGCAAGCGCCATA 892

QY 191 tcggattgggaaacatacaaaagccatcagttgggcaacctgttcacacgagcgagcca 250
   |||
Db 893 TCGGATTGGGAAAAATACAAGCCATAGTTGGGCAACCTGATGATTCACAGCGGCCCA 952

QY 251 ttaaggaataatcgtgctacattgtccgcttccgatacgggacgaagtcacatccc 310
   |||
Db 953 TTAAGGAATAATCGTGCTACATTGTCCGCTTTCCGATCAGCGGACGAAGTCATTCCC 1012

QY 311 ccttcgaaacacatcgtctacatccgattctgtatgaagccggtgtcccgattgaagcat 370
   |||
Db 1013 CCTTCGAACACATGCTCTCATTCGATTCTGATGAAAGCCGCTGATCCGTTGACGGAT 1072
```


QY 851 tcaagaactatgcgcgaagccatccgcgattggcgactgccaaaaccatgcgcgac 910
|||||
Db 3950 TCAAGACTATGCGCGACAGCATCCGGATTGGCGAGTCCAAAACCCATGCCGCAC 4009
QY 911 aagcgatagaagccgctcagcaaatctttaagcgagtcaccccgtoaaaggattggag 970
|||||
Db 4010 AAGCATAGAAAGCCGTCAGCATATCTTATGGCAGCCATCCCATCAAGAGGATTGGAG 4069
QY 971 ctgtcggggaagaataagcgcttggcgcgatcaagggacacatctctgtaaacggttcgaga 1030
|||||
Db 4070 CTGTTCCGGGAAATACGGCTTTGGCGGCATCAAGGCACTCTATCAACCGCTCGAGA 4129
QY 1031 tggcgagatcgcatttgcggaagggaatccgcgcgtcagcagaatlttgcgagtcgag 1090
|||||
Db 4130 TGGCGCGATCGCATTTGCCGAAGGGAATCCGCCCTCAGCCGACATTTTGGCGATCGG 4189
QY 1091 catacgcaaatatcccgctcccttaccatctccggaatatccgttcaacttggagcagc 1150
|||||
Db 4190 CATACGCCAAATACCCGCTTACCATCCGGAATATCCGTTCAAACTTGGAGCAGC 4249
QY 1151 gttaacgcaaaagaatacatcctcctcaacacgttgcgcgcgttaaacggaagaatgtga 1210
|||||
Db 4250 GTTACGGCAAGAAATACACCTCTCAACCGTGCCTCAAAAGGCAAAATGTCA 4309
QY 1211 aacttgcgaacaaacgcacccggaagacccaagtgccgttttgaagtggttcgga 1270
|||||
Db 4310 AACTGGCAGACCAACGCCGCGAAGACAGGCGTACCGTTTGAAGGATTTCGA 4369
QY 1271 atttggaaaagcgttaaatatcgatagc 1299
|||||
Db 4370 ATTTTGAAGACGCTGAATATGTACG 4398

RESULT 10
LOCUS AX239715 4425 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 17 from Patent WO0164920.
ACCESSION AX239715
VERSION AX239715.1 GI:15797369
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 4425)
AUTHORS Arico,M.B., Comanducci,M., Galeotti,C., Masiagnani,V., Giuliani,M.M.
and Piza,M.
TITLE Hybrid expression of neiserial proteins
JOURNAL Patent: WO 0164920-A 17 07-SEP-2001;
Chilton Spa (IT)
FEATURES
source location/Qualifiers
1..4425
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="delta983-ORF46.1"
BASE COUNT 1165 a 1259 c 1219 g 782 t
ORIGIN

Query Match 70.4%; Score 1176.2; DB 6; Length 4425;
Best Local Similarity 97.3%; Pred. No. 1.8e-266;
Matches 1196; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 71 cctcagatttgcgaacgatcttltatccgcgaggttctcagccgtcagcatttcgaac 130
|||||
Db 3170 CTCACGATTTGGCAAGATCTTTATCCGGCAGGTTCTCGACCGTCAAGCATTTGGAAC 3229
QY 131 ccgacgggaataatcacctatctgcgcagcaggggggaacttgcgagcgacggtcata 190
|||||
Db 3230 CCACGGGAATACACCTATTTCGGCAGCAGGGGGAACCTTCCGCGCCAGGGGCATTA 3289
QY 191 tccgatttggaaatacaagcagcatcagtttggggaacctgtctcagcagcgagcgaca 250
|||||
Db 3290 TCGGATTGGGAAAAATATCAAAAGCATCATGTTGGCAACCTGATGATTCAACAGCGGCCA 3349

QY 251 ttaagaagaatcgcgctacatcttccgctttccagatcaaggcaggaatccattcc 310
|||||
Db 3350 TTAAGAAATATCGCTATCATGTCGGCTTTCCGATTCACGGCGACGAAGTCCATTTCC 3409
QY 311 ccttcgacaacatlgctccatctccgatctcgtatgaagccggtatgtccgtttcagcat 370
|||||
Db 3410 CCTTGACAAACATGGCTCATTCGATTTCTGATGAAGCCGGTAGTCCCGTTGACGGAT 3469
QY 371 tcaagccttaccgcgcatccatcttgggaaggaatacgaacacatcccgcgagctatgag 430
|||||
Db 3470 TTAGCCTTTACCGCATATTCATTTGGAGCGATACGAACACCATCCGCGACGCTATGACG 3529
QY 431 ggcacagggcgcgagctatcccgctcccaaggcgcgagggatataatacgtcagaca 490
|||||
Db 3530 GGCACAGGGGGGGCTATCCGCTCCCAAGGCGGAGGATATATACAGTACGACA 3589
QY 491 taaaggcgttgcgcaaaatataccgcctcaacctgacgcgaacccgcagcagcagacaac 550
|||||
Db 3590 TAAAGGCGTTGGCCAAATATCCGCTCAACCTGACCGACGCAACCGCAGCACGACAC 3649
QY 551 ggccttgcgacggttccacaataaccggtatgcttgcgcaaggaaggtaggcagcgat 610
|||||
Db 3650 GCGTTGCCAGCGCTTCCCAATGCCGCTAGTATGCTGACGCAAGAGTATGCGCAGCGAT 3709
QY 611 tcaaacgcgcaccccgatataagcccgagcttggacagatggggaatggcgcggaagctt 670
|||||
Db 3710 TCAAAAGCCGACCCGATACGCCCGAGCTGGACAATGCGGCAATGCCCCCGAAGCCT 3769
QY 671 tcaacgcgcatcgcagatcgttcaaaaacatacatcgcgcgcgcgcgcgcgcgcgcgcgc 730
|||||
Db 3770 TCAACGGCACCTGCAGATATCGTTAAAAACATATCCGCGCGGCGAGAAATTTGCGGCG 3829
QY 731 caggcgatgcgcttgcaggtatataagcgaaggtcccaacatctgctatgcagcgcttgg 790
|||||
Db 3830 CAGGGGATCCCGTGCAGGGGATAGGGAAGGCTCAAAACATTGCTGATGCACGCGTTGG 3889
QY 791 gctcgttccacacgaagaatagtgcgcatcaacgatttggagataatggcgcaac 850
|||||
Db 3890 GTCCTCTTCCACCGCAAAACAGATGGCGGCATCAACGATTTGGCAGATATGGCGCAC 3949
QY 851 tcaagaactatgcgcgagcagccatccgcgatttggcgactgccaaaaccatgcgcgac 910
|||||
Db 3950 TCAAGACTATGCGCGACAGCATCCGGATTGGCGAGTCCAAAACCCATGCCGCAC 4009
QY 911 aagcgatagaagccgctcagcaaatctttaagcgagtcaccccgtoaaaggattggag 970
|||||
Db 4010 AAGCATAGAAAGCCGTCAGCATATCTTATGGCAGCCATCCCATCAAGAGGATTGGAG 4069
QY 971 ctgtcggggaagaataagcgcttggcgcgatcaagggacacatctctgtaaacggttcgaga 1030
|||||
Db 4070 CTGTTCCGGGAAATACGGCTTTGGCGGCATCAAGGCACTCTATCAACCGCTCGAGA 4129
QY 1031 tggcgagatcgcatttgcggaagggaatccgcgcgtcagcagaatlttgcgagtcgag 1090
|||||
Db 4130 TGGCGCGATCGCATTTGCCGAAGGGAATCCGCCCTCAGCCGACATTTTGGCGATCGG 4189
QY 1091 catacgcaaatatcccgctcccttaccatctccggaatatccgttcaacttggagcagc 1150
|||||
Db 4190 CATACGCCAAATACCCGCTTACCATCCGGAATATCCGTTCAAACTTGGAGCAGC 4249
QY 1151 gttaacgcaaaagaatacatcctcctcaacacgttgcgcgcgttcaaacggaagaatgtga 1210
|||||
Db 4250 GTTACGGCAAGAAATACACCTCTCAACCGTGCCTCAAAAGGCAAAATGTCA 4309
QY 1211 aacttgcgaacaaacgcacccggaagacccaagtgccgttttgaagtggttcgga 1270
|||||
Db 4310 AACTGGCAGACCAACGCCGCGAAGACAGGCGTACCGTTTGAAGGATTTCGA 4369
QY 1271 atttggaaaagcgttaaatatcgatagc 1299
|||||
Db 4370 ATTTTGAAGACGCTGAATATGTACG 4398

RESULT 11
AX236447
LOCUS AX236447 2019 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 140 from Patent WO0164922.
ACCESSION AX236447
VERSION AX236447.1 GI:15796045
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 2019)
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
Guiliani,M.M. and Pizzi,M.C.
TITLE Heterologous expression of neisserial proteins
JOURNAL Patent: WO 0164922-A 140 07-SEP-2001;
Chiron Spa (IT)
FEATURES
Source Location/Qualifiers
1..2019
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="ORF46.1-741"
BASE COUNT 555 a 571 c 531 g 362 t
ORIGIN

Query Match 70.4%; Score 1175.6; DB 6; Length 2019;
Best Local Similarity 97.2%; Pred. No. 2.2e-296;
Matches 1196; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 73 tcagattggcaaaagatcttcttaccgagcaggtctcgaaccgtcgaattcgaaacc 132
DB 4 TCAGATTGGCAAAAGATCTTTATCCGGCAGGTTCTCGACGTCAGCATTTGCAACC 63

QY 133 gacgggaataacacattcttgccagcaggggggaacttgcgcagcagcaggtcatc 192
DB 64 GACGGGAATACCACTTATTCGGCAGCAGGGGGAACCTTGCCGAGCGCAAGCCATATC 123

QY 193 ggaattggaaacatacaaacacatcagttgggaacctgttaccagcagcagccatt 252
DB 124 GGATTGGGAAAAATACAAAGCATCATGTTGGGCAACCTGATGATTAACAGGGGCCATT 183

QY 253 aaaggaatcgcgtacattgctcgcgtcttccgaatcagcagcagaatcattcccc 312
DB 184 AAAGGAATATCGGCTACATTTGCCGCTTTCCGATCAGCGGACGAAATCCATCCCC 243

QY 313 ttgcgaacacatgctcattccgattctgatgaagccggttagtccgtttgcgattc 372
DB 244 TTGCAACAACATGCTCATTCGATTCTGATGAAGCCGCTAGTCCGTTGACGGATT 303

QY 373 agccttaccgatacatgaggagcagatacgaaccattcccgacgagcgtatgaagg 432
DB 304 AGCCTTACCGCATTCATTGGGACGATACGAACACCATCCGCCGACGCTTATGACGG 363

QY 433 ccacagggcggcgtatccgcgtccccaaggcgcgagagatataatacagctacgacata 492
DB 364 CCACAGGGGGGGGCTATCCGCTCCCAAGGGGCGAGGATATATACACTACGACATA 423

QY 493 aaagcgcttgcccaataatccgctcaactcgtaccgacgaacacgcgacggaacagg 552
DB 424 AAAGCGCTTGCCCAAAATATCCGCTCAACCTGACCGCAACCGCAACCGCAACCGG 483

QY 553 ctgtgcgaacctttcacaaataccggtagtatgctgacgcaagaaggaataggcgaagtc 612
DB 484 CTTGCGCGACCGTTTCCACAAATGCGGTAGTATGCTGACGCAAGGAGTAGCGAGATT 543

QY 613 aaacgcgcaaccgatacagccccgagctgagacagatcgggcaatgcccgcgaagcttc 672
DB 544 AAACGCGCAACCGGATACAGCCCCGAGCTGAGACAGATCGGCAATGCGCGCAAGCCTTC 603

QY 673 aacgcgacatgcagatatacgtcaaaaacatcatcgcgcgcgcgagagaagaattgtcgcgca 732
DB 604 AACGCGACTGCAGATATTCGTTAAACATCATCGGCGCGGAGAGAAATTTGCGGCA 663

QY 733 ggcgatgcgtgcaagggtataagcgaagctcaaacattgtctgtatgcaagcgttgggt 792
DB 664 GGCATGCCGTCGCGAGGGGATAGGGAAGCTCAAAACATTTGGGTGTCATGCACGCGTTGGGT 723

QY 793 ctgcttccaccgaaacaagatgctgcgcatcaacgatttggcagatataggcgaactc 852
DB 724 CTGCTTCCACCGAAACACATGATGCGCGCATCAACGATTTGCGAGATATGCGCGCAACGC 783

QY 853 aaagactatgcgcaagcagccatccgattgggagatccaaacccaatgcccagaca 912
DB 784 AAAGACTATGCGCGAGCGCATCCGCTTGGGAGTCCAAAACCCCATCCGACACA 843

QY 913 ggcataaagcgcgtcagcaatatacttaccgagcagtcaccccgctcaaaaggattgagct 972
DB 844 GGCATAGAGCGCGTCAGCAATATCTTTATGGAGCGCATCCCATCAAGAGGATGGAGCT 903

QY 973 gttcggggaataatcagcttggcgagcatcagcagacatctctcaagcgcgtcgcagatg 1032
DB 904 GTTCGGGGAATAATACGGCTTGGGCGCATCACGGGCAATCTTATCAAGCGGTGCGAGATG 963

QY 1033 ggcgagatcgcgtatgcgcaaaagggaatccgcgcgtcagcgaacatttgcgatgcgca 1092
DB 964 GCGCGATCGCATTCGCGAAAGGGAATCCGCGTCAGCGAACAATTTGCCATGGCGCA 1023

QY 1093 tacgccaataaccggtcccttaccattcccgaaatattcgltcaaaacttggagagcgt 1152
DB 1024 TACGCCAAATACCGGTCCCTTACCATTCGCCGAATATCCGTTCAAACTTGGAGCAGCGT 1083

QY 1153 taagcgaagaagaacataactctctcaacgcgtgcccgcgttaaaaggaaagtgtgca 1212
DB 1084 TACGGCAAAAGAAACATACCTCTCAACCGTGGCGCGTCAAAACGCAAAATGTCAAA 1143

QY 1213 ctggcaaaccaagccacccgcaagacagcgaagtgcgtttgagcgttaaaaggtttccgaat 1272
DB 1144 CTGGCAGACCAACGCCACCCGGAAGCAGCGGTACCGTTGACGATGAAGGTTCCGAA 1203

QY 1273 ttgaaaaagacgttaaaatacagatcagaga 1302
DB 1204 TTTGGAAGCACGTGAATATGATACGGGA 1233

RESULT 12
AX239733
LOCUS AX239733 2019 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 35 from Patent WO0164920.
ACCESSION AX239733
VERSION AX239733.1 GI:15797377
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 2019)
AUTHORS Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Guiliani,M.M.
and Pizzi,M.C.
TITLE Hybrid expression of neisserial proteins
JOURNAL Patent: WO 0164920-A 35 07-SEP-2001;
Chiron Spa (IT)
FEATURES
Source Location/Qualifiers
1..2019
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="ORF46.1-741"
BASE COUNT 555 a 571 c 531 g 362 t
ORIGIN

Query Match 70.4%; Score 1175.6; DB 6; Length 2019;
Best Local Similarity 97.2%; Pred. No. 2.2e-296;
Matches 1196; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 73 tcagattggcaaaagatcttcttaccgagcaggtctcgaaccgtcgaattcgaaacc 132
|||||

Db 4 TCAGATTGGCAAAAGATCTTTATCCGGCAGTTCTGACCGTCAGATTTGAAACC 63
OY 133 gacggaaatccacccctattggcagcagggggaacttgcgagcgccaggtcatatc 192
Db 64 GACGGGAATATCCACTATTCGGCAGCAGGGGGAATCTGCCAGCGCACGGCCATATC 123
OY 193 ggaattggaaacatacaagaacgatcagttggaacactgtctcaacagcggcgccatc 252
Db 124 GGATTGGGAAAAAATACAAAGCATCAGTTGGGCAACCTGATGATTAACAGGGGGCCATT 183
OY 253 aaaggaatatcgctacatctgtccgctttccgatcacgggcaagatccatccccc 312
Db 184 AAAGGAATATTCGGCTACATTGTCCGCTTTTCGATTCACGGGCGACGAAGCCATTCCCC 243
OY 313 ttgacaacacatgctcaacttcgcattcgtatgaagccggatgctccgttgcagatc 372
Db 244 TTGACACACCTGCTCTCATTTCCGATTTCTGATGAAGCCGGTAGTCCCGTTGACGGATTT 303
OY 373 agccttaccgcatccatcttggaacgatatcagacacacatcccgccagcgctatgaagg 432
Db 304 AGCCTTACCGCATTCATTTGGAGCGATACGAACACATCCCGCCAGCGCTATGACGGG 363
OY 433 ccacagggcgcgcgctatcccgctcccaaggcgcgaggatataatacagtaagacata 492
Db 364 CCACAGGGCGCGGCTATCCCGCTCCCAAAGCGCGAGGATATATACAGCTACGACATA 423
OY 493 aaagcgctgcccacaaatccacgctccctcaacctgacacgaacacgcgacccggacagg 552
Db 424 AAAGGCTTGGCCAAAATATCCGCTCTCAACCTGACCGACACCGCAGCCGACACAGG 483
OY 553 ctgtgcagccgtttccaaataacgaagtagtctgacgcaagtagtagcgacgacatc 612
Db 484 CTGCGCCACCGTTTCCAAATGCGGAGTAGTATGCTGACGCCAAGAGTAGGCGACGATTCC 543
OY 613 aaacggcgaccccgatccagcccgagcttgacagatatcggaatgcccgcgaagcttc 672
Db 544 AAAGCGCCACCGCATATCAGCCCGAGCTGGACATCGGGGAATGCCCGCAAGGCTTTT 603
OY 673 aacgacactgcagatcagtcatacaaaacatccatcgcgcgaggaagaaattgctcgcgca 732
Db 604 AAGGCGACTCGAGATTCGTAAAAACATCATCGCGCGGCGAGGAATTTGTCGGCGCA 663
OY 733 ggcgtagcgctgtagggatataagcgaagctcaaacattgctgttatgcaagcttggt 792
Db 664 GGGGATGCCCTGCAAGGCAATGAAGGAAGCTCAAAACATTGCTGTCATGACGCGCTTGGGT 723
OY 793 ctgcttccacggaaacaaagattgcgcgcatccagatlltggcgatbtgycgacatc 852
Db 724 CTGCTTTCCACCGAAACAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAATTC 783
OY 853 aaagactatgcccagcagcagccatccgcatgttggaagtcacaaaccccaatgcccagaa 912
Db 784 AAAGACTATGCGCGAGCAGCATCCGGGATTTGGGCAATGCCAAACCCCATGCGCGACAA 843
OY 913 ggaatgaagacgctcagcaatatcttaagcgagtcacatcccgctcaaaaggaattgagct 972
Db 844 GGGATGAAACCCGTCACATATCTTATGGCAGCATCCCATCAAAAGGAGTTGAGACT 903
OY 973 gttcgggaaataagcgttggggcgcatccagacatccctcgtcaagaggtgcagagatg 1032
Db 904 GTTCGGGGAAATATACGCTTGGGCGCATCAAGGACATCTCATGAAGGGGTGCGCAGATG 963
OY 1033 ggcgaagatcgcatctgcgaagaaggaaatccgcgctcagcgaacatlltccgatgagga 1092
Db 964 GGGCGCATCGCATTTGCCGAAGGAATCCGCGTCACGACGAATTTTCCGATGGGGCA 1023
OY 1093 taagccaataaccgctcccttaccatcccgaaatalcogltcaaatcttgagcagcgt 1152
Db 1024 TACGCCAAATATCCGCTTACCATTTCCGAAATATCCGTTCAAACTTGGAGCAGCGT 1083
OY 1153 taaggaagaagaacatcacctcctcaacgctgcgcgctgaagaaggaagatgtaaa 1212
Db 1084 TACGCCAAAGAAACATCACCTCTCAACCGTGCAGCGCTCAAAAGCGCAAAATGTCAAA 1143

OY 1213 ctggcaacaacagccaccgcgaagaagtgccgtttgaagtgtaagggttccgaat 1272
Db 1144 CTGGCAGACCAACGCCACCCGAAGACAGGCGCTTGTACGCTTAAAGGCTTCCGAAAT 1203
OY 1273 ttgaaagaagacgttaataatagatagaga 1302
Db 1204 TTTGACAACACGCTGAATATGATACGGCA 1233

RESULT 13
AX236451 2256 bp DNA linear PAT 26-SEP-2001
LOCUS
DEFINITION Sequence 144 from Patent WO0164922.
ACCESSION AX236451
VERSION AX236451.1 GI:15796047
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE
1 (bases 1 to 2256)
AUTHORS
Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
Guilliani,M.M. and Pizze,M.C.
TITLE
Heterologous expression of neisserial proteins
JOURNAL
Chiron Spa (IT)
FEATURES
Source
1..2256
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="ORF46.1-961c"
BASE COUNT 698 a 608 c 545 g 405 t
ORIGIN

Query Match 70.4%; Score 1175.6; DB 6; Length 2256;
Best Local Similarity 97.2%; Pred. No. 2.2e-296;
Matches 1196; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

OY 73 tcagattggcgaagatcttltccggcagagttctgcagcgtcagcatlltgcgaacc 132
Db 4 TCAGATTGGCAAAAGATCTTTATCCGGCAGTTCTGACCGTCAGATTTGAAACC 63
OY 133 gacggaaatccacccatcttcggcagcagggggaacttgcgagcgccaggtcatatc 192
Db 64 GACGGGAATATCCACTATTCGGCAGCAGGGGGAATCTGCCAGCGCACGGCCATATC 123
OY 193 ggaattggaaacatacaagaacgatcagttggaacactgtctcaacagcggcgccatc 252
Db 124 GGATTGGGAAAAAATACAAAGCATCAGTTGGGCAACCTGATGATTAACAGGGGGCCATT 183
OY 253 aaaggaatatcgctacatctgtccgctttccgatcacgggcaagatccatccccc 312
Db 184 AAAGGAATATTCGGCTACATTGTCCGCTTTTCGATTCACGGGCGACGAAGTCATTTCCCC 243
OY 313 ttgacaacacatgctcaacttcgatttgatgaagccgtagtccgcttgcagatc 372
Db 244 TTGACACACCTGCTCTCATTTCCGATTTCTGATGAAGCCGGTAGTCCCGTTGACGGATTT 303
OY 373 agccttaccgcatccatcttggaacgatatcgaacacatcccgccagcgctatgacgg 432
Db 304 AGCCTTACCGCATTCATTTGGAGCGATACGAACACATCCCGCCAGCGCTATGACGGG 363
OY 433 ccacagggcgcggtatcccgctcccaaggcgcgaggatataatacagtaagacata 492
Db 364 CCACAGGGCGCGGCTATCCCGCTCCCAAAGCGCGAGGATATATACACTACGAATA 423
OY 493 aaagcgctgcccacaaatccacgctccctcaacctgacacgaacacgcgacccggacagg 552
Db 424 AAAGGCTTGGCCAAAATATCCGCTCTCAACCTGACCGACACCGCAGCCGACACAGG 483
OY 553 ctgtgcagcgtttccaaataacgtagtagtctgacgcaagtagtagcgacgacatc 612
Db 484 CTGCGCCACCGTTTCCAAATGCGGAGTAGTATGCTGACGCCAAGAGTAGGCGACGATTCC 543

QY	Db	484	CTTGGCAGCCGTTTCCACAAATGCCGGTAGTATGCTGACGCAAGAGATGAGCGACGATTC	543
QY	613	aaacgcgcacccccatatacaagcccccagcttgagacagatcgggcaatgcccgcgaagcttc	672	
Db	544	AAACGGCCACCCGATACACCCCGAGCTGGACGATCGGCAATGCGCGGAAGCCTTC	603	
QY	673	aaagcagctcagatatacgttcaaaaatacatacggcgcgccgaagaatattgtcgggca	732	
Db	604	AACGGCACTGCAAGTATCGTTTAAAMCATCATCGGCCCGGAGGAGAAATTTGTGGGCA	663	
QY	733	ggcagatccgctgcagggtataaagcgaagcttcaacaactctgttatgcacgcttgggt	792	
Db	664	GGCGATCCCGTCGACGGGCATTAAGCGAAGCGCTCAAAATTTGTGTCATGACGAGCTTGGGT	723	
QY	793	ctgctttccaccggaaaaaaagaatgagcgacatcaacgattggcagatataggcgaatc	852	
Db	724	CTGCTTTTCCACCGTAAACAGATGGCCGCTATCAACCATTTGGCAGATATGGCCCAATC	783	
QY	853	aaagacatgscgaagaagacatcccgatggatgggacgttccaaaaccccaatgscgacaa	912	
Db	784	AAAGACTATGCGCAGCAGCCATCCGGAATTGGCAGTCCAAAACCCCAATGCCGCACAA	843	
QY	913	ggcataagaagccgtcagcaaatatctttacggcagtcatacccccgtcaaaaggtatggact	972	
Db	844	GGCATAGAGAGCGGTACAGAAATATCTTTATGGCAGACCAATCCCATCAAAAGGATTTGGAGCT	903	
QY	973	gttcggggaaaaatacgcgcttggtggcgacatcaagcgaacatctcttcaagcggttcgagatg	1032	
Db	904	GTTCCGGGGAAAAATACGGCTTTGGGGGCTCACAGGCACATCTTATCAACCGGTCCAGATG	963	
QY	1033	ggcgagatcgcgatctgcccgaagggaatccgcgcgtcagcagaacaatttgcgcagtgcgca	1092	
Db	964	GGCGCGATTCGCATTCGCCAAAGGGAATCCGCCGTCAAGCAATTTTGGCGATGCGGCA	1023	
QY	1093	tacgcgaataacccgcttcccttaacoattcccggaatatcgttcaaaccttggagcagcgt	1152	
Db	1024	TACGCCAAATACCCGTCCTCTTACCATTTCCGGAATATCTGTTCAAACTTGAGAGAGCGT	1083	
QY	1153	tacggcaagaagaacaatcactctctcaaacgcttgcgcgcgttcaaaaggaagaatgtgaa	1212	
Db	1084	TACGGCAAGAAACATCACTCTCTCAACCTGCGCGCGCTCAAAAGGCAAAATGTCAAA	1143	
QY	1213	cttggcaaaaacaacgcgccaccggaagccaagaatgtccgttttgaacgttaaaaggtttccgaat	1272	
Db	1144	CTGGCAGACCAACGCCACCCCAAGACAGCGCTTACCGTTTGACGGGTAAAGGGTTTCCGAAT	1203	
QY	1273	tttggaaaaagcgttaataatagatagaga	1302	
Db	1204	TTTGAAGAGCACGTGAATATGATACGGGA	1233	
RESULT 14				
LOCUS	AX239737	2256 bp	DNA	linear
DEFINITION	Sequence 39 from Patent WO0164920.			
ACCESSION	AX239737			
VERSION	AX239737.1	GI:15797379		
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
BASE COUNT				

[illegible]

[illegible]

Db	CCACAGGCGGCGCGCTATCCCGCTCCCAAGGCGCAGGGATATATACAGCTACGACATA	423
QY	aaagcggtttgccccaaaatctcgccctcaacttaccgaacacgcgcacccgaacagc	552
Db	424 AAGGCGGTTTCCCAAAATATCCGCTTCAACCTGACGACACCCGACACCCGACAA	483
QY	ctgttcgacggtttccacaatacgcgttaqatlatgtctgacgaagagtaagcgacgttc	612
Db	484 CTGGCGGACCGTTTCCACAAATGCGGCTAGATGCTCAGCAAGAGATAGGCGACGATTC	543
QY	613 aaacgycgcaacccgatacagccccgcagcttgcagacatcgggcaatgcgcgcgaagcttc	672
Db	544 AAACGCGCACCCGATACAGCCCGAGCTGGACAGATCGGGCAATGCCCGCAAGCCTTC	603
QY	673 aacgcgacgtgcagatatcgtcaaaaatcatatcgctggcgcgaggaataattgtcgcgca	732
Db	604 AACGCGACTGTGCAATATTCGTTAAAAACATATATGCGCGCGACGCAAAATTTGCGGCGCA	663
QY	733 ggcgcatagcggttgcgaaggtataagcgaaggtcacaacatctgcgttatacgcgcttgggt	792
Db	664 GCGGATGCGGCTGAGCGGCTAAGCGCAAGGCTCAAAACATTGCTGTCTGACGCGCTTGGGT	723
QY	793 ctgcgtttccacccgaaaacgaagtatggcgcatcacaagatttgcagatatgtgcgaatcc	852
Db	724 CTGCTTTCCACCGAAAAACAAGATGGCGGCATCAACGATTTGGCAATATGGCGCAATC	783
QY	853 aaagactatgcgcgcgcgcacacacacgcgcgttgcgcgtcacaaccccccaatgcgcacaa	912
Db	784 AAAGACTATATGCGCGACGACGCCATTCGCCGATTTGGGAGTGTCCAAACCCCATGCGGACAA	843
QY	913 ggcataagaagccgctgaagcaatatccttacgcgaatcatcccccgtcaaaaggattggaact	972
Db	844 GGCATATGAAGCCGCTCAGCAATATCTTTATATGCGACGCATCCCATCAAAAGGATTTGGAGCT	903
QY	973 gttcgcgggaataatacgcttggcggcgtacaaggaacatccctgtaagcggtccgaagatg	1032
Db	904 GTTTCGGGAAAAATACGGCTTGGCGGCAATCAGGCACATCTATCAAGCGGTGCGCAGATG	963
QY	1033 ggcgcgatacgcatttgcgcgaagaaggaatccgcgcgttcgaaggaacatttgcgcgtgcgca	1092
Db	964 GCGCGGATTCGCAATTCGCCGAAAGGAATGCCGCGTACGAGCAAAATTTTCCGATGCGGCA	1023
QY	1093 taagcgcnaataacgcgtcccccctaccattccgcgaataatccgttcaacttgcgcgcgcgt	1152
Db	1024 TACGCGCAATATCCCTCCCTTACCATTTCCCGAAATATTCGTTCAAACTGTGAGCAGGCT	1083
QY	1153 taagcgaagaagaataatacctcctccaacccgtgcgcgcgttcaaaaggaagaattgtgaaa	1212
Db	1084 TACGCGCAAAAGAAACATCACTCTCAACCGTGCCTCAAAAGGCAAAAAATGTCAAA	1143
QY	1213 ctgcgaacaacacgcgcacccgcgaagcacaagatgycggtttgacggttaaaaggtttccgcat	1272
Db	1144 CTGGGAGAACCAAGCGCACCCGACGAGACAGCGGATTCGTTGACGATTAAGGCTTTCGGAAT	1203
QY	1273 ttgaaagaagcgttaaatatcogatacgsaga	1302
Db	1204 TTTGAGAAGCACGTTGAATATGATACGGGA	1233
RESULT 16		
LOCUS	AX239735	
DEFINITION	Sequence 37 from Patent W00164920.	
ACCESSION	AX239735	
VERSION	AX239735.1 GI:15797378	
KEYWORDS		
SOURCE		
ORGANISM	synthetic construct.	
REFERENCE	synthetic construct	
AUTHORS	1 (bases 1 to 2421)	
	Arico,M.B., Comanducci,M., Galeotti,C., Mesignani,V., Giuliani,M.M.	
	and Pizzo,M.	

TITLE Hybrid expression of neisserial proteins
JOURNAL Patent: WO 0164920-A 37 07-SEP-2001;
Chiron Spa (IT)
FEATURES
source
1. .2421
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="ORF46.1-961"
BASE COUNT 730 a 659 c 591 g 441 t
ORIGIN

Query Match 70.4%: Score 1175.6; DB 6; Length 2421;
Best Local Similarity 97.2%: Pred. No. 2.2e-296;
Matches 1196; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 73 tcaaatltggcaaacgattcttltatccggcaggttctcgacgttcgaacatttcgaacc 132
Db 4 TCAGATTGGCAAGCATCTTTATCGGGAGGTTCTCGACGCTCAGCATTTCCGAACCC 63

QY 133 gaagggaataccacctatctcgacagcagggggaacttgcagagcagcgatcatc 192
Db 64 GACGGGAATACCACTATTCGGCGAGCGGGGAATTGCGAGCGCAAGCGCATATTC 123

QY 193 ggaattggaaaacatataaagccatcagtttggcaacctgtlcatccagcaggcgccatt 252
Db 124 GGATTGGGAAAATACAAAGCCATCAGTTGGCAACCTGATGATTAACAGCGGCGCATTT 183

QY 253 aaaggaaatccggtacattgtccgcttltccgaatccagggcaagtlccattccccc 312
Db 184 AAAGGAATATCGGCTACATCTCCGCTTTCCGATCACGGGACAGAAATCCATTCCTCCC 243

QY 313 tttgcaaacacctgtctacattccgattctgtatgaagccggatgtccgtttgacgattc 372
Db 244 TTGCAACAACCTCCCTACATTCGATTCGTATGAAGCCGGATGTCCGTTGCGGATTT 303

QY 373 agccttaccgcatccattgtgacgagatacgaacacatcccgccgacggtatgaagcgg 432
Db 304 AGCCTTACCGCATCCATTGGGAGGATACGAACCAATCCGCCGACGCGCTATGACGGG 363

QY 433 ccaagggcgcggtatcccgctcccaaggcgagggatataatacagctacgacata 492
Db 364 CCACGGGCGGGGCTATCCCGCTCCCAAAAGCGCGAGGATTAATACAGCTACGACATA 423

QY 493 aaagcggttgcacaaatatactgcctcaacctgacagcaaacccgacgacgacaacg 552
Db 424 AAAGCGTTGGCCAAAATATCCGCTCAACCTGACGACAAACCGGACGACGACAAACGG 483

QY 553 ctgtgcgaccttccacaataccggtatgtatgtcgtacgcaaggagtaaggcagcgtatc 612
Db 484 CTTGCCGACCGTTTCCACAATGCCGTACTATGCTGACGCAAGGAGTAAAGCGAGGATTC 543

QY 613 aaagcgcaaccgcatatacagcccgagctggacaagatcgggcaatcgcgcggaagcttc 672
Db 544 AAACCGGCAACCCGATACGCCCGAGTGAGACATCGGGCAATCCCGCGAAGCCCTTC 603

QY 673 aacggcactgagatatacgtcaaaaacatcaltcgcgcgcgcaagaagaatltgcgcgca 732
Db 604 AAGGCACTGCGATATCTTTAAACATCATTCGGCGGCGGAGAGAAATTTGCGGCGCA 663

QY 733 ggcagatgcggtcagggatlaagcgaagctcaaacattgtcttatacgcggtctggt 792
Db 664 GGCGATGCGGTCCAGGGCTAAAGCGAAGGCTCAAACTGCTGTCAACGCGGCTTGGGT 723

QY 793 cgtgttccacgaaaaaagaatgycgcgcatcaacagatttggcagatgtgagcgaactc 852
Db 724 CTGCTTTCCACCGAAACAAAGATGGCGGCATCAACGATTTGGCAGATATGGCGCACTC 783

QY 853 aaagactatgcgcgcaagccatccgcatltggcagtlccaaaaaccacatgcccgcacaa 912
Db 784 AAAGACTATGCGGACAGCGCATCCGGAATGGGAGTCCAAAAACCCCAATGCGGCACAA 843

QY 913 ggcataagaagcggtcagcaatactttaagcgagtcattcccgctcaaggatgtgagct 972

Db 844 GGCATACAGCCGTCACCAATATCTTTATGCGACGATCCCATCAAGGATTTGGAGCT 903

QY 973 gtccggggaataatcagcttggcggtatccagcaactctctgtcaaacggttcgaagt 1032
Db 904 GTTCGGGGAATAATCGGCTTGGCGGATCGGCGACATCTTATCAACGGGTCGAGATG 963

QY 1033 ggcgagatcgcattgcccgaaggggaatccgcgtlcaagcgaactttgcgagtgcgca 1092
Db 964 GGCGGATCGCATTTGCCGAAAGGGAATCCCGCGTACAGCAAAATTTTCCGATCGGCA 1023

QY 1093 taagcaaataccggtcccttaccattcccgaaatcgtlccaaacttggagcagcgt 1152
Db 1024 TACGCCAATATCCGTCCTTACCATTCGGAATATCCGTTCAACTTGGAGCAGCGT 1083

QY 1153 taagcaagaagaacatcactctctcaacggttgcgcggttaaaaggaagatltgaa 1212
Db 1084 TACGGCAAGAAACATCACCTCTCAACCGTCCGCGTCAAAAGCGCAAAATGTCAAA 1143

QY 1213 ctggcaacaacagccacccgaagaccagatgcccgttgaacggtlaaagggtlccgaat 1272
Db 1144 CTGGCAAGCCAACGCCACCGCAAGACAGCGGTACGTTGACGTTAAAGGTTTCGAAAT 1203

QY 1273 ttgaaaaagacgtlaaatacagatacagaga 1302
Db 1204 TTGAGAAAGCACGTGAATATGATACGCGA 1233

RESULT 17
AX236465 2304 bp DNA Linear PAT 26-SEP-2001
LOCUS
DEFINITION Sequence 158 from Patent WO0164922.
ACCESSION AX236465
VERSION AX236465.1 GI:15796054
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
1 (bases 1 to 2304)
REFERENCE
AUTHORS
Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
Guillani,M.M. and Pizzi,M.C.
TITLE
Heterologous expression of neisserial proteins
JOURNAL
Patent: WO 0164922-A 158 07-SEP-2001;
Chiron Spa (IT)
FEATURES
source
1. .2304
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="961cl-ORF46.1"
BASE COUNT 710 a 619 c 555 g 420 t
ORIGIN

Query Match 70.3%: Score 1174.4; DB 6; Length 2304;
Best Local Similarity 97.1%: Pred. No. 4.6e-296;
Matches 1196; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 73 tcaaatltggcaaacgattcttltatccggcaggttctcgacgttcgaacatttcgaacc 132
Db 1069 TCAGATTGGCAAGCATCTTTATCGGGAGGTTCTCGACGCTCAGCATTTCCGAACCC 1128

QY 133 gaagggaataccacctatctcgacagcagggggaacttgcagagcagcgatcatc 192
Db 1129 GACGGGAATACCACTATTCGGCGAGCGGGGAATTGCGAGCGCAAGCGCATATTC 1188

QY 193 ggaattggaaaacatataaagccatcagtttggcaacctgtlcatccagcaggcgccatt 252
Db 1189 GGATTGGGAAAATACAAAGCCATCAGTTGGGGAACCTGATGATTAACAGGCGGCGCATTC 1248

QY 253 aaaggaaatccggtacattgtccgcttltccgaatccagggcaagtgatccattccccc 312
Db 1249 AAAGGAATATCGGCTACATTCGCTTTTCCGATCAGCGGCGACGAAGTCATTCCTCCC 1308

QY 313 ttccgaacacatgcctacatccgattctgatgaagccggtlgtccggttcagcgaattc 372
|||||
Db 1309 ttccgaacacacatgcctacatccgattctgatgaagccggtlgtccggttcagcgaattc 1368
QY 373 agccttaccgcatcatctggagcagatagcaaacacaccccgccgagcgtctatgcggg 422
|||||
Db 1369 agccttaccgcatcatctggagcagatagcaaacacaccccgccgagcgtctatgcggg 1428
QY 433 ccaagggcgcgcgctatccgctcccaaaaggcggaagatatatacagctacagacata 492
|||||
Db 1429 ccacagggcgcgcgctatccgctcccaaaaggcggaagatatatacagctacagacata 1488
QY 493 aaagcgcttgcggcaaaatccgctcccaaaaggcggaagatatatacagctacagacata 552
|||||
Db 1489 aaagcgcttgcggcaaaatccgctcccaaaaggcggaagatatatacagctacagacata 1548
QY 553 ctgttcgacgcttcccaaatccggtatgctgtagcgaaggaatagcagcagattc 612
|||||
Db 1549 ctgttcgacgcttcccaaatccggtatgctgtagcgaaggaatagcagcagattc 1608
QY 613 aaacgcgccacccgcatcagcccgagctggaacagatcgggaatgcccgcgaagcttc 672
|||||
Db 1609 aaacgcgccacccgcatcagcccgagctggaacagatcgggaatgcccgcgaagcttc 1668
QY 673 aaagcgacgtgagatattctcaaaaaacatccatccgcgcggaaggaatgttcgcgcga 732
|||||
Db 1669 aaagcgacgtgagatattctcaaaaaacatccatccgcgcggaaggaatgttcgcgcga 1728
QY 733 ggcgcatcgctgagcaggtatagcgaagcctcaaacattgctgttatgacgagcttggt 792
|||||
Db 1729 ggcgcatcgctgagcaggtatagcgaagcctcaaacattgctgttatgacgagcttggt 1788
QY 793 ctgcttccacgcaaaaaaagaatgagcgcatcgaagatlttgacagataltgagcgaactc 852
|||||
Db 1789 ctgcttccacgcaaaaaaagaatgagcgcatcgaagatlttgacagataltgagcgaactc 1848
QY 853 aaagacatitcgcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 912
|||||
Db 1849 aaagacatitcgcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1908
QY 913 ggcataagaagcgctcagcaatattctttagcgaagtcaccccgctcaaaaggaattgagct 972
|||||
Db 1909 ggcataagaagcgctcagcaatattctttagcgaagtcaccccgctcaaaaggaattgagct 1968
QY 973 gttcggggaataatcagcttggcgcgcatcagcagcagcagcagcagcagcagcagcagcagc 1032
|||||
Db 1969 gttcggggaataatcagcttggcgcgcatcagcagcagcagcagcagcagcagcagcagcagc 2028
QY 1033 ggcgagatcgcattgcccgaaggaagaaatccgcttcaagcgaatlttgccgattgcgca 1092
|||||
Db 2029 ggcgagatcgcattgcccgaaggaagaaatccgcttcaagcgaatlttgccgattgcgca 2088
QY 1093 taagcgaataatccgctcccttaccatcccggaataltcagcttcaaacattggaagcagct 1152
|||||
Db 2089 taagcgaataatccgctcccttaccatcccggaataltcagcttcaaacattggaagcagct 2148
QY 1153 taaggaaaaaaaacatcactcctcaacgctgcccgcgtcaaaaggaagaatgtgaaa 1212
|||||
Db 2149 taaggaaaaaaaacatcactcctcaacgctgcccgcgtcaaaaggaagaatgtgaaa 2208
QY 1213 ctggcaaaaaaacgcaagcaggaagcaagctgagcttggcaggtlaaagggttccgatt 1272
|||||
Db 2209 ctggcaaaaaaacgcaagcaggaagcaagctgagcttggcaggtlaaagggttccgatt 2268
QY 1273 tttagaaaaagcgtaaaaatagcagtagagaat 1304
|||||
Db 2269 tttagaaaaagcgtaaaaatagcagtagagaat 2300

RESULT 18
AX239751
LOCUS AX239751 2304 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 53 from Patent WO0164920.

ACCESSION AX239751
VERSION AX239751.1 GI:15797386
KEYWORDS
SOURCE
ORGANISM synthetic construct.
ARTIFACT synthetic construct.
REFERENCE 1 (bases 1 to 2304)
AUTHORS Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
TITLE Hybrid expression of neisserial proteins
JOURNAL Patent: WO 0164920-A 53 07-SEP-2001;
Chiron Spa (IT)
FEATURES
source Location/Qualifiers
1..2304
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="961cI-ORF6.1"
BASE COUNT 710 a 619 c 555 g 420 t
ORIGIN
Query Match 70.3%; Score 1174.4; DB 6; Length 2304;
Best Local Similarity 97.1%; Pred. NO. 4.6e-296;
Matches 1196; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 73 tcaagatttgcaaaagcattcttattccgagcaggttctcgacgttcaagcatttcgaacc 132
|||||
Db 1069 tcaagatttgcaaaagcattcttattccgagcaggttctcgacgttcaagcatttcgaacc 1128
QY 133 gacgggaataataccactatccgacgagcagggggaacttgcagcagcagcagcagcagcagc 192
|||||
Db 1129 gacgggaataataccactatccgacgagcagggggaacttgcagcagcagcagcagcagcagc 1188
QY 193 ggaatttggaacacatacaagccatcagcttgaggcaacctgttcacatcccgcaagcgagcatt 252
|||||
Db 1189 ggaatttggaacacatacaagccatcagcttgaggcaacctgttcacatcccgcaagcgagcatt 1248
QY 253 aaaggaataatcggctacatgttcgcttccgatacagcagcagcagcagcagcagcagcagc 312
|||||
Db 1249 aaaggaataatcggctacatgttcgcttccgatacagcagcagcagcagcagcagcagcagc 1308
QY 313 tttagaacaacatgctcctcaacttccgattctgtagaagcggtagtcccggttagcagattc 372
|||||
Db 1309 tttagaacaacatgctcctcaacttccgattctgtagaagcggtagtcccggttagcagattc 1368
QY 373 agccttaccgcatcacttggagcagatagcaaacacatcccgcgcgagcagcagcagcagcagc 432
|||||
Db 1369 agccttaccgcatcacttggagcagatagcaaacacatcccgcgcgagcagcagcagcagcagc 1428
QY 433 ccaagggcgcgcgctatccgctcccaaaaggcgcgaggaatatalatacagctacagacata 492
|||||
Db 1429 ccaagggcgcgcgctatccgctcccaaaaggcgcgaggaatatalatacagctacagacata 1488
QY 493 aaaggggtggcccaaatatccgcttcccaaacctgacagcagcaacccgcaagcagcagcagc 552
|||||
Db 1489 aaaggggtggcccaaatatccgcttcccaaacctgacagcagcaacccgcaagcagcagcagc 1548
QY 553 ctgttcgacgcttcccaaatatccgcttcccaaacctgacagcagcaacccgcaagcagcagcagc 612
|||||
Db 1549 ctgttcgacgcttcccaaatatccgcttcccaaacctgacagcagcaacccgcaagcagcagcagc 1608
QY 613 aaacgcgccacccgcatcagcccgagctggaacagatcgggaatgtccgcgcgaagcttc 672
|||||
Db 1609 aaacgcgccacccgcatcagcccgagctggaacagatcgggaatgtccgcgcgaagcttc 1668
QY 673 aaagcgacgtgagatattctcaaaaaacatccatccgagcagcagcagcagcagcagcagc 732
|||||
Db 1669 aaagcgacgtgagatattctcaaaaaacatccatccgagcagcagcagcagcagcagcagc 1728
QY 733 ggcgcatcgctgagcaggtatagcgaagcctcaaacattgctgttatgacgagcttggt 792
|||||
Db 1729 ggcgcatcgctgagcaggtatagcgaagcctcaaacattgctgttatgacgagcttggt 1788

QY 793 ctgcttccaccggaacaaagatgycgcgatcaacgatttgcagatgatalgycgcaactc 852
|||||
Db 1789 CTGCTTTCACCCGAAACAAAGATGGCGCGATCAAGCATTTGGCAGATATGGCGCACTC 1848
QY 853 aaagactatgycgcagcagccatccgcgatttgcgcaatcccaaaaccccaatgycgcgcaaa 912
|||||
Db 1849 AAAGACTATGCCCGACGACGATCCGGATTTGGCAGTCCAAAACCCCAATGGCGCACAA 1908
QY 913 ggcataagaacgcgtcagcaatattcttaacgcaatccatcccgctcaaaagagattggaact 972
|||||
Db 1909 GGCATAGAACCGCTGACGATATCTTTATGGCAGCAGCATCCCATCAAAAGGATTTGGAGCT 1968
QY 973 gttcggggaanaatacggccttgcgcgcatcaaggcaacatccctgtcaagcgcgttcgagatg 1032
|||||
Db 1969 GTTCGGGGAANAATTCGGCTTGGCGCGCATCAGGCGACATCTCATCAACCGGTCCGACAGATG 2028
QY 1033 ggcgagatgcgcattgcggaaggggaatccgcgcgtcaacgcaaatatttgcgcgttcgca 1092
|||||
Db 2029 GCGCGATCGCATTTGGCGAAGGGAATCCCGCTGACGCGAATTTTGGCGATCGCGCA 2088
QY 1093 taacgcaaatatcccgctcccttaccattcccgaaatatccgttcaaatcttgcagcagcgt 1152
|||||
Db 2089 TAGCGCAAAATACCGTCCCTTACCATTCGCGAAATATCCGTTCAAACTTGAGAGCAGCGT 2148
QY 1153 taacgcaaaagaataatcactctctcaacgctgcgcgcgttaaaagaaagattgaa 1212
|||||
Db 2149 TAGCGCAAAAGAAATCATCTCTCAACCGTCCGCGTCAACGCGAAGGGAATGTCAAA 2208
QY 1213 ctgcgaacaaacgacgcacccggaagcaaaagtgcgcttgcagctaaaggttccgaat 1272
|||||
Db 2209 CTGGCAGACCAACGCCACCGCAAGACAGCGGTACGTTTGACGGTAAAGGTTTCCGAAT 2268
QY 1273 ttgaaaaaagacgtaaatatcagatacgaagaat 1304
|||||
Db 2269 TTTGAGAAGCACGTGAATATGATATCATCAACT 2300

RESULT 19
AX236459 2256 bp DNA linear PAT 26-SEP-2001
LOCUS AX236459
DEFINITION Sequence 152 from Patent WO0164922.
ACCESSION AX236459
VERSION AX236459.1 GI:15796051
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 2256)
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
TITLE Heterologous expression of neisserial proteins
JOURNAL Patent: WO 0164922-A 152 07-SEP-2001;
Chilton Spa (IT)
FEATURES
source 1..2256
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="961c-ORF46.1"
BASE COUNT 698 a 608 c 545 g 405 t
ORIGIN

Query Match 70.3%; Score 1174.2; DB 6; Length 2256;
Best Local Similarity 97.3%; Pred. No. 5,1e-296;
Matches 1194; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 73 taagatttgcgaacagatcttattatccgcgcaagttctcgacgctcagcatcttgaacc 132
|||||
Db 1003 TCAGATTTGGCAAAAGCATCTTTATTCGCGACAGTTCTGCAACGCTCAGCATTTGAAACC 1062
QY 133 gaacggaataatcacactattctgcgcagcagggggaacttgcgcgagcgcaagcgtcatatc 192
|||||
Db 1063 GACGGGAAATACCACTATTTCGCGACAGGGGGAACTTGCGGAGCGAGCGGCGCATATTC 1122

QY 193 gatttgggaacatacaaaagcattcagtttgcgaacctgttcatccagcagcgccatt 252
|||||
Db 1123 GGATTGGGAAATAATACAAACCATTCAGTTGGGCAACTGTGATTTCAACAGCGGCCATT 1182
QY 253 aaagaataatcgcgtacatctgtccgcttccgatcacaggcaggaagtcattccccc 312
|||||
Db 1183 AAAGGAATATCGGTGATACATTTGTCGCTTTCCGATCACGGGACGAAGTTCATTCCCCC 1242
QY 313 ttgcacaacatgctctcaatctccgatcttgatgaagcggtagtccggttgaaggatc 372
|||||
Db 1243 TTCGACAACCATGCTTCACATTCGATTCGATTCGATTAACCCGGTATCCCGTACGAGATT 1302
QY 373 agcctttaccgcaatccatttgcgagatacgaacaacatcccgcgcaagcgtcatgacgag 432
|||||
Db 1303 AGCTTTACCGCATCTCAATTTGGAGCGATACCAACACCATCCCGCGAGCGGTATGACGGG 1362
QY 433 ccacagggcgcggtctatcccgctcccaagcgcgcgaggaatataatagctacgacata 492
|||||
Db 1363 CCAAGGCGCGCGCTATCCCGCTCCCAAGCGCGGAGGATATATACAGCTACGACATA 1422
QY 493 aaagcgcttgcgaataatctcgcgtcaacctgaacgcaacccgacgacgagcaacg 552
|||||
Db 1423 AAAGCGCTTGCCTCAAAATATCCGCTCAACCTGACCGCAACCGACGACGACGACACGG 1482
QY 553 ctgtgcacgcttcccaaatatccggttagta tgcgtacgcgaaggaagtaggcgcgattc 612
|||||
Db 1483 CTTCGCCACCGTTTCCAAATATGCGGTATGCTGACGCAAGGAGTATGCGACGATTC 1542
QY 613 aaacgcgcacccgatacagccccgcgagcttgacagatcgtgggaatgcccgcgaacgttc 672
|||||
Db 1543 AAAGCGCGCACCCCATATACGCCCGAGCTGGAACAATGGGGGAATGCCGCCAAGCGTTTC 1602
QY 673 aacgcgactgagatatcgtcaaaaacataatcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 732
|||||
Db 1603 AACGCGACTGAGATATCTTAAAAACATCATCGGCGCGCGGAGGAATTTGTCGGCGCA 1662
QY 733 ggcgattgcgctgcaggtgataagcgaagctcaaacattgctgttatgcagcgcttgg 792
|||||
Db 1663 GCGCATGCGCGTAGAGGCGATTAAGCGAAGGCTCAAACTTGTCTGATGACGCGCTTGGGT 1722
QY 793 ctgcttccaccggaacaaagatgycgcgatcaacgatttgcagatattgcgcgaac 852
|||||
Db 1723 CTGCTTTCACCGGAACAAAGATGGCGCATCAACGATTTGGCAGATATTTGGCGCAACTC 1782
QY 853 aaagactatgycgcagcagccatccgcgatttgcgagcttcccaaaaccccaatgcgcgca 912
|||||
Db 1783 AAAGCATATGCCCGGACGACGCGCATCCGCGATTTGGCGCATCAAAACCCCAATGCCACAA 1842
QY 913 ggcataagaagccgttagcaatattcttaacgcaatctcccgctcaaaagagattggaact 972
|||||
Db 1843 GGCATAGAAAGCGTACGCAATCTTTATGGCAGCGCATCCCATCAAAAGGATTTGGAGCT 1902
QY 973 gttcggggaanaatacggccttgcgcgcatcaacgcaacatccctgtcaagcgcgttcgagatg 1032
|||||
Db 1903 GTTCGGGGAATAATACGCTTTGGCGCGCATACCGGCAACATCTTCAAGCGGTGCGAGATG 1962
QY 1033 ggcgagatcgcattgcgcgaaggaatccgcgcgttcagcgcagcaatttgcgcgttcgca 1092
|||||
Db 1963 GCGCGGATCGCATTTGCCGAAAGGGAATCCGCGTCAAGGACACATTTTGGCGATTCGCGCA 2022
QY 1093 taacgcaaatatcccgctcccttaccattcccggaatatccggttcaaatcttgcagcagcgt 1152
|||||
Db 2023 TAGCGCAAAATACCGCTCCCTTACCATTCGCGAATAATTCGTTCAAACTTTGGAGCAGCGT 2082
QY 1153 taacgcaaaagaataatcactctctcaacgcttgcgcgcgttcaaaaggaagattgaa 1212
|||||
Db 2083 TAGCGCAAAAGAAATCATCTCTCAACGCTCCGCGCTCAAAAGGCAAAATATGTCAAA 2142
QY 1213 ctgcgaacaaacgcgcacccggaagcaagctgcgcttgcggttgcggttaaggttccgaat 1272
|||||
Db 2143 CTGGCAGACCAACGCGCACCGCGAAGACAGCGGTACCGTTTACGCTTAAGGTTTCCGAAT 2202

QY 1273 ttgaaaaagacgtataacacg 1299
||||| 11 ||||| 11111
Db 2203 TTTGAGAACGACGTGAATATGATACG 2229

RESULT 20
AX239745 2256 bp DNA linear PAT 26-SEP-2001
LOCUS AX239745
DEFINITION Sequence 47 from Patent WO0164920.
ACCESSION AX239745
VERSION AX239745.1 GI:15797383
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
1 (bases 1 to 2256)
REFERENCE
Arlino, M.B., Comanducci, M., Galeotti, C., Massignani, V., Giuliani, M.M.
and Pizzi, M.
Hybrid expression of neisserial proteins
Patent: WO 0164920-A 47 07-SEP-2001;
Chiron Spa (IT)
FEATURES
source
Location/Qualifiers
1..2256
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="961c-ORF46.1"
BASE COUNT 698 a 608 c 545 g 405 t
ORIGIN

Query Match 70.3%; Score 1174.2; DB 6; Length 2256;
Best Local Similarity 97.3%; Pred. No. 5.1e-296;
Matches 1194; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 73 ttgatttgcaaacagatctctttatccgagcggtcttcgacgctgaacattcgacc 132
Db 1003 ttgatttgcaaacagatctctttatccgagcggtcttcgacgctgaacattcgacc 1062

QY 133 gacggaataacacacattcgcgacgagggggaacttcgcgagcgagcgatc 192
Db 1063 GACGGGAATATACACTATTCGGCAGCAGGGGGAATTCGGCAGCGCCATATC 1122

QY 139 gatttggaataacacacattcgcgacgagggggaacttcgcgagcgagcgatc 252
Db 1123 GATTTGGAATAATACACTATTCGGCAGCAGGGGGAATTCGGCAGCGCCATATC 1182

QY 253 aaagaaatcgcgtacattcgcgctttccgatacaggggaacgaattccccc 312
Db 1183 AAAGAAATATCGGTACATTTCGCGCTTTTCGATCAGGGGAGAGATCCATTCGCC 1242

QY 313 ttgcgaacacatgctcattccgattctgataagccggtagcttcgattgacgattc 372
Db 1243 TTTGCAACCAATGCTTCATTCGGCAGGATACGAACACATTCGCCGCGATATGACGG 1302

QY 373 agccttaccgacatccattgagcagatagcaaacacatcccgccgagcgtatgacgg 432
Db 1303 ACCCTTACCGCATCATTTGGGAGGATACGAACACATTCGCCGCGATATGACGG 1362

QY 433 ccacagcgcgcgctcattccgctcccaagcgcgagggatataacagctacgacata 492
Db 1363 CCACAGCGCGGCGCTATCCCGTCCCAAGGCGCAGGAGATATATACAGCTACGACAT 1422

QY 493 aaagcgctgcacaaataatcgcctcaactgaacgaacacgagcagcccgagacag 552
Db 1423 AAAGCGCTTGCCCAAAATATTCGCTCAACCTACCGACACACGCGACACGCAACGG 1482

QY 553 ctgtgacgctttccacatacggtagtatagtcgtacgcaagagtagcgacgattc 612
Db 1483 CTGTGCGACCGTTTCCACATGCGGTAGTATGCTACGCAAGAGTAGGCGACGATTC 1542

QY 613 aaacgagcagcagatagacgcccgcgagctggagcagatggcagctccgacgaattc 672
Db 1543 AAACGCGCACCCGATACAGCCCCGAGCTGAGCAGATCGGGCAATCGCCGCGACCTTC 1602

QY 673 aacgacgtcagatatactcctcaaaaacatcatcgcgcgaggaanaatttcgagcga 732
Db 1603 AACGCACTGCAATATTCGTATAAAACATCATTCGCGCGGCGAGAAATATTCGCGGCA 1662

QY 733 ggcgattgcgctgaggtataagcgaagcctcaaacattgctgtatgcaagcgttggt 792
Db 1663 GCGGATGCGGTGCGAGGGCATAAACGAAGGCTCAAAACATTCGTCTCATGACAGGCTTGGGT 1722

QY 793 ctgcttccacgcaaaaacagatgcgcgcatcaacagatttgcagatagcgcaactc 852
Db 1723 CTGCTTCCACGCAAAACAAAGATGGCGCGCATCAAGATTTGGCAATATGGCGCACTC 1782

QY 853 aaagactatgcgcagcagccatccgagattggcagctccaaaaaccccaatgcgcaca 912
Db 1783 AAAGACTATGCCGCAACAGCATCCCGATTTGGCGAGTCCCAAAACCCCAATGCCGCA 1842

QY 913 ggcataagacgctcagcaatatactttagcgagatcccgctcaagagattgagct 972
Db 1843 GGCATGAAAGCCCTCAGCANTATCTTTATGCGACGCCATCCCATCAAAAGGATTTGAGCT 1902

QY 973 gtcgaggaaatacagcttggtgagcagatcaagcagatcgtcgaagcggtcgagatg 1032
Db 1903 GTTCGGGGAATATACGCTTGGCGGCATACACGACATCTCTATCAAGCGGTGCGAGATG 1962

QY 1033 ggcgagatgcatttcgcaagaaagaaatccgcgctgaagcaaatcttcgagtcgga 1092
Db 1963 GCGCGATGCGATTGCGCAAGGAAATCCGCGTACGCAACATTTTCCGATGCGGCA 2022

QY 1093 taagcaaatlaccgctcccttaccatctccgaataatccgtcaaacattgagcaagct 1152
Db 2023 TACGCCAAATATACCGCTTCCCTTACATTCGCCGAATATCCGTTCAAACTTGAGCGCGG 2082

QY 1153 ttgcgcaaaaacatacctctcctcaacgctgcgcgctcaagcaagaaatgctga 1212
Db 2083 TACGGCAAAAGAAATACATCTCTCCTACCGGTCGCGCGCTCAACGCAAAATGTCAAA 2142

QY 1213 ctggcaaaaacacgcaacccgaaagacaaagtcgcttgtagcgttaagggcttcgaa 1272
Db 2143 CTGGCAACCAACGCAACCGCAAGACAGGCGTACCGTTGACGATTAAGGCTTCCGAAT 2202

QY 1273 ttgaaaaagacgtataacacg 1299
Db 2203 TTTGAGAACGACGTGAATATGATACG 2229

RESULT 21
AX236453 2421 bp DNA linear PAT 26-SEP-2001
LOCUS AX236453
DEFINITION Sequence 146 from Patent WO0164922.
ACCESSION AX236453
VERSION AX236453.1 GI:15796048
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
1 (bases 1 to 2421)
REFERENCE
Arlino, M.B., Comanducci, M.C., Galeotti, C.C., Massignani, V.C.,
Giuliani, M.M., and Pizzi, M.C.
Heterologous expression of neisserial proteins
Patent: WO 0164922-A 146 07-SEP-2001;
Chiron Spa (IT)
FEATURES
source
Location/Qualifiers
1..2421
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="961-ORF46.1"
BASE COUNT 730 a 659 c 591 g 441 t
ORIGIN

Query Match 70.3%; Score 1174.2; DB 6; Length 2421;
Best Local Similarity 97.3%; Pred. No. 5.2e-296;

Matches 1194; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
QY 73 tcaatttggcaaaagatcttcttccggcaggttctgcacggttaagatttcgaacc 132
    |||||||
Db 1168 TCAGATTTGGCAAAAGATCTTTTATCCGGAGGTTCTGACCGTACGATTTTCGAACCC 1227
QY 133 gacggaatccaccattctcgccagcaggggaaatttcgcagcgacggttcattc 192
    |||||||
Db 1228 GACGGGAATACACCTTTTCGGCAGCAGGGGGAATCTGCCAGGCGACGCCCATATC 1287
QY 193 ggaattggaaacatacaaaagccatcaglttggcaacctgttcaatcagaagcgccatc 252
    |||||||
Db 1288 GGATTTGGAAAAATACAAAGCATCATGTGGCAACCTGATGATCAACGCGCGCATTT 1347
QY 253 aaagaaatatcggttacctatcttcgcttttcgatacgggagcggaagttccatccccc 312
    |||||||
Db 1348 AAAGGAAATATCGCTACATTTGTCGCTTTCCGATCAGGGGACGACGAATTCCTCCCC 1407
QY 313 ttcgacaacacatgctccatctcgaatctgaatgaagccgttagtccgtttcgagatlc 372
    |||||||
Db 1408 TTTCGACAAACATCGCTCAATTCGATTCGATGGAAGCCGGTGTCCCGTTGACGGATT 1467
QY 373 agccttaaccgcatccatcttggagcgataagaaaccatccgcgcgacgtatgacgg 432
    |||||||
Db 1468 AGCCTTTACCGCATTCATTTGGGACGGATACGAACACATCCCGCCGCTATGACGGG 1527
QY 433 ccacagggcgggcggtatcccgctcccaaggcgagagatataataagctacgacata 492
    |||||||
Db 1528 CCACAGGGGCGGGCTATCCCGCTCCCAAGGCGGAGGATATATACACTTACGACATA 1587
QY 493 aaagcgcttgcacaaataatccgcctcaactgacgcgaacacgcgcgcgcgcgcgcgc 552
    |||||||
Db 1588 AAAGCGCTTGCCCAAAATATCCGCTCAACCTGACGACGAACCGCAGCACGGCAACGG 1647
QY 553 ctgtcgacacgttctccacaataccggttagtgaatgaagagatgaagcgacggttc 612
    |||||||
Db 1648 CTGTGCGACCGCTTTCCACATATCCCGTACTATGCTACGCAAGAGATGACGAGGATT 1707
QY 613 aaagcgcaacccgataacgcccgcagctgagacagatcggcgcaatgcgcgcgaagcttc 672
    |||||||
Db 1708 AAACCGCCACCCGATACAGCCCGAGCTGAGACAGATCGGCGAATGCCCGCAAGCCTTC 1767
QY 673 aacgacatgacagatatcgtlcaaaaaacatcaltcgycgcgcgcgcgcgcgcgcgcgc 732
    |||||||
Db 1768 AACGGACACTGAGATATCTTTAAACATCATGCGCGCGCAGAGAAATTTGTCGGCGCA 1827
QY 733 ggcgatgcgtgcaggggttaagcgaagctcaaacattgctgttatagcgcgcttggt 792
    |||||||
Db 1828 GCGGATGCGCTCAGGGCATTAAGCGAAGGCTCAAAATTCGTCTCATGACACGGCTTGGGT 1887
QY 793 ctgcttccacgcaaaaagaatgycgcgcatacaagatttggcagaatattggcgcaatc 852
    |||||||
Db 1888 CTGCTTTTCCACGGAANAACAGATGCGCGCATACCATTTGGCAGATATGCGCACTC 1947
QY 853 aaagactatgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 912
    |||||||
Db 1948 AAAGACTATGCGCGACAGCAGCATCGCGANTTGGGAGTCCAAAAACCCATGCGCGCAAA 2007
QY 913 ggcataagaagcgtcagcaatatcttlaagcgatcaatcccgctcaaaaggtattgagct 972
    |||||||
Db 2008 GGCATAGAAAGCCGTGACGATATCTTTATGGCAGCCATCCCATCAAAAGGATTTGGAGCT 2067
QY 973 gttcgagggaataatagcgtcttggcgcgacatacctcttcaagcggttcgaatg 1032
    |||||||
Db 2068 GTTCCGGGAAAATACGGCTTTGGCGGCATACCGCATCTTATCAACCGGTTCCCAATG 2127
QY 1033 ggcagatcgcattcgcgaagaagaaatccgcgctcagcgcaaatlttgcgatgcgca 1092
    |||||||
Db 2128 GCGCGGATGCAFTGCGGAAGGAAATCGCGCTCAGCGACAATTTTGGCGATGCGGCA 2187
QY 1093 taagccaataaccgctcccttaacgattccggaataatccggttcaaaacttggagacgct 1152
    |||||||
Db 2188 TACGCCAAATACCGTCCCTTACCATTTCCGAAATATCCGTTCAAACTTGACACGCGT 2247
```

```
QY 1153 taagcaaaagaacatcacctcctcaacggtgcgcgcgtcacaagcaagaatgtgaaa 1212
    |||||||
Db 2248 TACGGCAAAAGAAACATACCTCTCTCAACCGCTGCGCGCTCAACGCAAAATGTCAAA 2307
QY 1213 ctggcaaacacacgcgcacccgaagaacaaagtgcggtttgacggttaagggttccgaat 1272
    |||||||
Db 2308 CTGGCAGACCAACGCCACCCGGAAGACAGCGCTTGTGACGGTAAAGGTTTCCGAAT 2367
QY 1273 ttgaaaaagacgttaaaataagatagc 1299
    |||||||
Db 2368 TTTGAGAAAGCAGCTGAATATGATAGC 2394

RESULT 22
AX239739
LOCUS AX239739 2421 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 41 from Patent WO0164920.
ACCESSION AX239739
VERSION AX239739.1 GI:15797380
KEYWORDS
SOURCE
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 2421)
AUTHORS Arico,M.B., Comanducci,M., Galeotti,C., Maignani,V., Giuliani,M.M.
and Pizzi,M.
TITLE Hybrid expression of neisserial proteins
JOURNAL Chiron Spa (IT)
FEATURES
source 1..2421
Location/Qualifiers
BASE COUNT 730 a 659 c 591 g 441 t
ORIGIN

Query Match 70.3%; Score 1174.2; DB 6; Length 2421;
Best Local Similarity 97.3%; Pred. No. 5.2e-296;
Matches 1194; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 73 tcaatttggcaaaagatcttcttccggcaggttctgcacggttaagatttcgaacc 132
    |||||||
Db 1168 TCAGATTTGGCAAAAGATCTTTTATCCGGAGGTTCTGACCGTACGATTTTCGAACCC 1227
QY 133 gacggaatccaccattctcgccagcaggggaaatttcgcagcgacggttcattc 192
    |||||||
Db 1228 GACGGGAATACACCTTTTCGGCAGCAGGGGGAATCTGCCAGGCGACGCCCATATC 1287
QY 1273 ggaattggaaacatacaaaagccatcaglttggcaacctgttcaatcagaagcgccatc 252
    |||||||
Db 1288 GGATTTGGAAAAATACAAAGCATCATGTGGCAACCTGATGATCAACGCGCGCATTT 1347
QY 253 aaagaaatatcggttacctatcttcgcttttcgatacgggagcggaagttccatccccc 312
    |||||||
Db 1348 AAAGGAAATATCGCTACATTTGTCGCTTTCCGATCAGGGGACGACGAATTCCTCCCC 1407
QY 313 ttcgacaacacatgctccatctcgaatctgaatgaagccgttagtccgtttcgagatlc 372
    |||||||
Db 1408 TTTCGACAAACATCGCTCAATTCGATTCGATGGAAGCCGGTGTCCCGTTGACGGATT 1467
QY 373 agccttaaccgcatccatcttggagcgataagaaaccatccgcgcgacgtatgacgg 432
    |||||||
Db 1468 AGCCTTTACCGCATTCATTTGGGACGGATACGAACACATCCCGCCGCTATGACGGG 1527
QY 433 ccacagggcgggcggtatcccgctcccaaggcgagagatataataagctacgacata 492
    |||||||
Db 1528 CCACAGGGGCGGGCTATCCCGCTCCCAAGGCGGAGGATATATACACTTACGACATA 1587
QY 493 aaagcgcttgcacaaataatccgcctcaactgacgcgaacacgcgcgcgcgcgcgcgc 552
    |||||||
```

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
LOCUS	Sequence 459 from Patent WO924578.	A96426	A96426							
DEFINITION	Sequence 459 from Patent WO924578.	A96426	A96426							
ACCESSION	Sequence 459 from Patent WO924578.	A96426	A96426							
VERSION	Sequence 459 from Patent WO924578.	A96426	A96426							
KEYWORDS	Sequence 459 from Patent WO924578.	A96426	A96426							
SOURCE	Sequence 459 from Patent WO924578.	A96426	A96426							
ORGANISM	Sequence 459 from Patent WO924578.	A96426	A96426							
REFERENCE	Sequence 459 from Patent WO924578.	A96426	A96426							
AUTHORS	Sequence 459 from Patent WO924578.	A96426	A96426							
TITLE	Sequence 459 from Patent WO924578.	A96426	A96426							
JOURNAL	Sequence 459 from Patent WO924578.	A96426	A96426							
FEATURES	Sequence 459 from Patent WO924578.	A96426	A96426							
source	Sequence 459 from Patent WO924578.	A96426	A96426							

[illegible]

TITLE	Complete genome sequence of <i>Neisseria meningitidis</i> serogroup B strain MC58	gene	complement(2550. .2936)
JOURNAL	Science 287 (5459), 1809-1815 (2000)	CDS	/gene="NMB0367"
MEDLINE	20175755		complement(2550. .2936)
PUBMED	10710307		/gene="NMB0367"
REFERENCE	2 (bases 1 to 11700)	gene	/note="hypothetical protein; identified by Glimmer2; putative"
AUTHORS	Tetzelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwin, M.L., Deboy, R., Peterson, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E., Clifton, H., Clark, E.B., Cotton, M.D., Ulterback, T.R., Khouri, H., Qian, H., Vamathevan, J., Gill, J., Scarlato, V., Maignani, V., Pizzi, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R., Rapunoli, R. and Venter, J.C.	CDS	/codon_start=1
TITLE	Direct Submission	gene	/transl_table=1
JOURNAL	Submitted (17-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr. Rockville, MD 20850, USA	CDS	/product="hypothetical protein"
FEATURES	Location/Qualifiers	gene	/protein_id="AAF40810.1"
source	1.11700	gene	/db_xref="GI:7225590"
	/organism="Neisseria meningitidis MC58"	CDS	/translation="MWDVAVKTPPELPFTSVGIFRGADITQYKNILETFWYPPDEFRTVEYSPSNLISVKKKNIISICYOELPMGVNIIGLNEDEFQLPHYKSYGVDKYVLSNESYPTVYVEFEIGQAWMAKGIWIIANGKDYSTEPYDD"
	/strain="MC58"	gene	complement(3444. .3833)
	/db_xref="taxon:122586"	CDS	/gene="NMB0369"
	/note="serogroup: B"	gene	complement(3444. .3833)
	201. .974	CDS	/gene="NMB0369"
	/gene="NMB0364"	gene	complement(3444. .3833)
	201. .974	CDS	/gene="NMB0369"
	/note="similar to GB:I06299 PID:293962 percent identity: 99.61; identified by sequence similarity; putative"	gene	/note="hypothetical protein; identified by Glimmer2; putative"
	/codon_start=1	gene	/codon_start=1
	/transl_table=11	gene	/transl_table=11
	/product="FPC operon protein"	gene	/product="hypothetical protein"
	/protein_id="AAF40807.1"	CDS	/protein_id="AAF40811.1"
	/db_xref="GI:7225587"	gene	/db_xref="GI:7225591"
	/translation="MRPVATTIYOLFIFLGSVYMTSCPEVNEQTSFNNDEPMTFEHTYEDPGTKWVPIGYLARYONNATKWLSDPGDAYSINLIEISVYKKKIDQGVLEPYQONKAHFIQFLRDLSDYVDIVIKKDACLSLITMGERRLLITGVKKPSPARPEVVEYEDKRHPENPFHEFYIKGENPAILTHRNKRIQTEEDSYSTSVGCLNGFTVRYPPIREKQQLTQOELVGYHOVEQLVQSFVNNSSK"	CDS	complement(3965. .4342)
	991. .1266	gene	/gene="NMB0370"
	/gene="NMB0365"	gene	complement(3965. .4342)
	/note="This region contains a match to at least one other gene that is not full length, and is not the result of a sequencing artifact; similar to GB:I06299 PID:293963 SP:P5127 percent identity: 98.72; identified by sequence similarity; putative"	CDS	/gene="NMB0370"
	complement(1353. .2486)	gene	/note="hypothetical protein; identified by Glimmer2; putative"
	/gene="NMB0366"	gene	/codon_start=1
	complement(1353. .2486)	gene	/transl_table=11
	/gene="NMB0366"	gene	/product="hypothetical protein"
	/note="hypothetical protein; identified by Glimmer2; putative"	CDS	/protein_id="AAF40813.1"
	/codon_start=1	gene	/db_xref="GI:7225593"
	/transl_table=11	gene	/translation="MCFKRPFRRIIPCFKEYDENSFIGKWYDGVMDDEEYKLENDLIEVRKYPYPMIDPRVYVIGITIDFLAVQNMKLFELIKASTWLPKSKYINERTFRVMRYLFTDVIDMDKWFYPIQHSKGLRRKV"
	/product="hypothetical protein"	gene	complement(4777. .5454)
	/protein_id="AAF40808.1"	CDS	/gene="NMB0372"
	/db_xref="GI:7225588"	gene	complement(4777. .5454)
	/translation="MFKRKPVLLSPFALVFAWLGIGIAYEINRFLSPATVPKNPNAVAKLARLFRNADRAVYLVKESIRTEENLAGTVDDGPLQSEKDYIALIRLSRLKERAKHVEHQEGKEVMDYHIGEGGLAVSLQSRSPEAFVNAEYLYRNDPFSVNVGGTVHGENYETTGRRVWVDPDGSVDAAGRGKIGDVEYHLCGYOMQOVYLAQYRDVANDOKWDPFKESNRITASRNSFYONMRDLMRGMKANSLVGYDADGLPKYVNSPDGKKRQSGFEYTLKNGNLFIAOSYVALKADCTYADMQTYHAOQTYLDDGRIVREKQGRLLPDPPLNLENLEKVRVYAEAAARSGGRDLSH"	CDS	complement(4777. .5454)

[illegible]

	AL62758	AL157959
ACCESSION	AL62758	AL157959
SOURCE	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;	Neisseria meningitidis 22491.
KEYWORDS	Neisseria meningitidis 22491	
ORGANISM	Neisseria	
REFERENCE	1 (bases 1 to 195767)	
AUTHORS	Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C., Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T., Davies,R.M., Davis,P., Devlin,K., Felkell,T., Hamlin,N., Holtiyd,S., Jagers,K., Leather,S., Moulse,S., Mungall,K., Quail,M.A., Rajandream,M.A., Rutherford,K.M., Stammers,M., Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.	
TITLE	Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491	
JOURNAL	Nature 404 (6777), 502-506 (2000)	
MEDLINE	20222556	
REFERENCE	2 (bases 1 to 195767)	
AUTHORS	Parkhill,J.	
TITLE	Direct Submmission	
JOURNAL	Submitted (30-MAR-2000) Submitted on behalf of the Neisseria sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk	
COMMENT	NOTES: Details of N. meningitidis sequencing at the Sanger Centre are available on the World Wide Web. (URL, http://www.sanger.ac.uk/projects/N.meningitidis/). Location/Qualifiers 1. 195767	
FEATURES	/organism="Neisseria meningitidis 22491" /strain="22491" /db_xref="taxon:122587"	
SOURCE	/note="serogroup: A" 80..952 /gene="NMA2045" 80..952 /gene="NMA2045" 80..952 /note="NMA2045", possible oxidoreductase, len: 290aa: similar to many eg. SW:P20692 (TYRA_BACSU) prephenate dehydrogenase from Bacillus subtilis (372 aa) fasta scores: E(): 2.8e-30, 39.4% identity in 284 aa overlap." /codon_start=1 /transl_table=1 /product="putative oxidoreductase" /protein_id="CAB85263.1" /db_xref="GI:7380673" /db_xref="SPTRMBL:Q9JY12" /translation="MPILNHAIQLGVGIGSEFVLDKRGCLVRPTVGIDTDNRNLERASLEGVIDRASVAIDADISIGADIVLIATPAVTAPALTALRPVLPEHTWISDGSSTSSVEAFPRCLPDRLRHRCIAAHPIAGSDRSGAQAOGFLPFHRKLITPHGESHDSGTALINLMHWAGAIEYITMDAORHDNVFAAVSHMPLTAAFYVHOILLDHDDGGEYLKFAPA TGFDLPRIIASGHAPVAMVDICLANKDSLLGLVGGIKGDVLAIDLTVADREARYRFEEAKTRDRWLNDGN" complement(547)..556) /note="Core DNA uptake sequence: gccgcctgaa" /label=DUS 960..969 /note="Core DNA uptake sequence: gccgcctgaa" /label=DUS 1009..2211 /note="Core DNA uptake sequence: gccgcctgaa" /label=DUS complement(1009..2211) /gene="NMA2046" complement(1009..2211) /gene="NMA2046" /note="NMA2046", hypothetical integral membrane protein, len: 400aa: similar to PR:p95556 (EMBL:Z73914) hypothetical protein from Pseudomonas stutzeri (396 aa) fasta scores: E(): 0, 35.0% identity in 391 aa overlap. Contains possible membrane-spanning hydrophobic regions." /codon_start=1 /transl_table=1 /product="hypothetical integral membrane protein" /protein_id="CAB85264.1"	


```

Db 79280 ACACCGCTTTCCGAGACAGATACGAGAACGCCCCCTTCGACAAACACGCCGCCG
QY 332 attcgaattctgtagaagccggtagtccttgtaacgattcagccttaccgcatcatt 391
Db 79340 ACAGCGCGAGCGAGAAAAGGCAACCTTACGAAAGCGCTTTACCGTATACGGGCTCACT
QY 392 ggaagcagatagacacacattcccgccgaagcgtatagcagcgcaagcgagcggtatc 451
Db 79400 GGGAGGACACGAAATCATCCCGCCGATGCTACGACGGCCCAAGGCGCGCATATTACC
QY 452 ccgctcccaaacgcgagggagatatatacagctacagacataaaagcggtgcccataa 511
Db 79460 CCAACCTAGCGGTGACGCGAGCAATACACCTATACGATCAACGCGACACAGCAGCA 79519
QY 512 tccgctcaactgacgcaaacgcaagcagcagcagcagcgtgtgacccgtttccaca 571
Db 79520 TCMAACTCAATCCGACGACACCGGACATCCGCAACGATATCCGACATTTACAGCA 79579
QY 572 ataccgtagtatgctgaagcaagtagtagcagcagattcaaacgcgcacccagatata 631
Db 79580 ACCTCGGGACCAATTTCTCCGACCGCGCCGATGAACCAAGAAAATGTTGAGCACA 79639
QY 632 gccccgaagctggaagaatcggaatcgccgaagcttcaacgagcctgcagatalcg 691
Db 79640 ATGCCAAGCTCGACCGCTGGGCGAAGCATGAGATTATCAACGCGCTCGCGCGCGCG 79699
QY 692 tcaaaacatcatcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 751
Db 79700 CGCTCAACCCCTTATATAGCGCGGCGGCAAGCCTTGGGCGATAGCGACATACTGTA 79759
QY 752 taagcgaagctcaaacatgctgtatgacagcgcttgctgcttccacgcgaataa 811
Db 79760 CGGGCTATGCGATAGAAAGCCGCAATGCGCAATGCGCCCTTGCCCGCGAGGGA 79819
QY 812 agatgagcgcatcaacagattggcagatatgagcgaactcaagaactatccgcagcag 871
Db 79820 AATTCCGCGCATCGCGGCGCTTGGGCGAGCGCTTGGAAAAATATACGCGCAAG 79879
QY 872 ccattcgagattggcagcttccaaaccccaatgcccagcagcagcagcagcagcagc 931
Db 79880 CCGTGGACCGGTGATACAGAAAACCCCAATGCCCGGAAACCTCGAAGCCCTGCTCA 79939
QY 932 atacttaccgagcagta 949
Db 79940 ACGTCTGCGGCTTGCCA 79957

RESULT 27
AE002420 10057 bp DNA linear BCT 25-MAY-2000
LOCUS Neisseria meningitidis serogroup B strain MC58 section 62 of 206 of
DEFINITION the complete genome.
ACCESSION AE002420 AE002098
VERSION AE002420.1 GI:7225876
KEYWORDS
SOURCE Neisseria meningitidis MC58.
ORGANISM Neisseria meningitidis MC58.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
1 (bases 1 to 10057)
Tetelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
Dodson,R.J., Nelson,W.C., Gwin,M.L., DeBoy,R., Peterson,J.D.,
Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
Cittone,H., Clark,E.B., Cotton,M.D., Uterback,T.R., Khouri,H.,
Qin,H., Yamahavan,J., Gill,J., Scarlato,V., Masignani,V.,
Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
Rappuoli,R. and Venter,J.C.
Complete genome sequence of Neisseria meningitidis serogroup B
strain MC58
Science 287 (5459), 1809-1815 (2000)

```

```

MEDLINE
PUBMED
REFERENCE
AUTHORS
20175755
10710307
2 (bases 1 to 10057)
Tetelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
Dodson,R.J., Nelson,W.C., Gwin,M.L., DeBoy,R., Peterson,J.D.,
Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
Cittone,H., Clark,E.B., Cotton,M.D., Uterback,T.R., Khouri,H.,
Qin,H., Yamahavan,J., Gill,J., Scarlato,V., Masignani,V.,
Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
Rappuoli,R. and Venter,J.C.
Direct Submission
Submitted (17-Mar-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
1..10057
/organism="Neisseria meningitidis MC58"
/strain="MC58"
/db_xref="taxon:122586"
/note="serogroup: B"
126..350
/gene="NMB0649"
126..350
/gene="NMB0649"
/gene="NMB0649"
/note="hypothetical protein; identified by glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA041070.1"
/db_xref="GI:7225877"
/translacion="MRLATKDLAEAIRKGGVRRKSSFTTDLRAIEGSESKIPDTYWH
HODTGRMOLIREGLHMDTGHIGMEANRGR"
323..742
/gene="NMB0650"
323..742
/gene="NMB0650"
/note="hypothetical protein; identified by glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA041071.1"
/db_xref="GI:7225878"
/translacion="MGSNEQKRYTMKRIKESDDEFAIKLFQSOSIDINEFLMTIE
OVARMDPIEDIPYIFDLADFGDIADINIVGVSSCRSKSKNMLTGIAFLRGID
VYDPPISKERAKLAKLKHREIYOKFOHFFVELPLP"
843..1232
/gene="NMB0651"
843..1232
/gene="NMB0651"
/note="hypothetical protein; identified by glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA041072.1"
/db_xref="GI:7225879"
/translacion="MKRIKESDDEFAIKLFQSOSIDINEFLMTIEOVARMDPIED
IPYIFDLADFGDIADINIVGVSSCRSKSKNMLTGIAFLRGIDVYDPPISKER
AKLAKLKHREIYOKFOHFFVELPLP"
1355..2296
/gene="NMB0652"
1355..2296
/gene="NMB0652"
/gene="NMB0652"
/note="similar to GP:4838554 percent identity: 79.35; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="mafa protein"
/protein_id="AA041073.1"
/db_xref="GI:7225880"

```

/translation="WKTLILLIPVLTACGTTIGTIPAGGGKRFVDEDELAASSRAA
YKMDLSLAKGRKALIVSYMGDDSGNI SGRYSIDLIRGYNHNPESAOY SYPA
VDTATTKSDALSSTTSTSLNAPPAALTKNSGRKERSAGLSVNGTDRNETLLA
NRPDVSFLTNLIQVFIYLRGIEVPEPADVDVETVADVEGTVSRSTELHLYNAETLK
AQRKLEFAVDRDSRKLITPRTKAYESQYOYOLMTGPYVSKTVKASDLMDFS
DTPYGGTTAQRNDRDFKONNKKPDVGNVIRRRGG"
2300. 3568
/gene="NMB0653"
2300. 3568
/gene="NMB0653"
/note="Maifb-related protein; identified by Oxford;
putative"
/codon_start=1
/transl_table=11
/product="Maifb-related protein"
/protein_id="AAF62314.1"
/db_xref="GI:7413436"
/translation="MKPLRLTNLACAVAAALILPALADLAQDPFITDNRQIH
YEPGKYHLFGDPRGSYSDRTKINVTIDYTHQNGNLLIQANINGTIGYHTRFSGHG
HEHPAPDNNAADSASEKGNVDGFTVYRLNMEGHEHHPADAYDGPNGNYPKPTGA
RDEYTYHVNSTANSIKLNPDTBSIRORI SPNYSNLGSNFSRADENARKMEHNAKL
DRGNSMEFINGVAGALNPFISAGEAVDQMOENPNAETVEALVNVLPRAKVVNL
KAAPGKAAVSGDFSDSKYKNTASRLSOSVDGEMEPFNVDFKASIGTCTIHDCQOK
HISGHRNYIEGKSTLNQINNPOLINGHSATYVYISKANGNPVDFGTPISDGKS
GLSTNCTHSGKNGVHIVPANPKTIKKVQ"
3571. 3888
/gene="NMB0654"
3571. 3888
/gene="NMB0654"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF41074.1"
/db_xref="GI:7225881"
/translation="MNILPSLIRVGNMIAIMLGTHSIRLITVDEGRRLKTKNVL
SREATREDHEDMEXLITELMSMCGEYFDEADPECTISNHSMEINQINGAVFERRELI
SOA"
4023. 4730
/gene="NMB0655"
4023. 4730
/gene="NMB0655"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF41075.1"
/db_xref="GI:7225882"
/translation="MYKTPADGYKALIRIRGCRVAFKDEASGKTGKVPYARRGNPYO
ETVYIEISDGIGNNQTLLISNKIHPETYSQGWIOAGRKKGDTLSESGAQTQYONITF
KQPLKATNLYIADWHITFYVKSQAELEGWVHNDCEPYDGNQRYKDSAYHGKNDNV
KSRAPTNQGAALDINSVOYKSTSPRRVGVKANNEIVLNTQTFNNSAAYHGHVRSW
ODLHTDOKNAALKKAGLD"
4756. 5085
/gene="NMB0656"
4756. 5085
/gene="NMB0656"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF41076.1"
/db_xref="GI:7225883"
/translation="MTDKSKTEKLISDDKOSYIDSIILMVENSKAYPEWPISEKIME
LSKNKDLDIAGSLTFCFGLHARLHNSIGDYDVYPIPLHKKQDPELQGRADLEBIS
LFLSENH"
5180. 5383
/gene="NMB0657"
5180. 5383
/gene="NMB0657"
/gene="NMB0657"

/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF41077.1"
/db_xref="GI:7225884"
/translation="MREFFKNPTLASSTAAPMSKROTAEKQTPSVLPPFDEYGSTP
MPPKPSKPSSTLPPPPKQSMOKR"
5211. 5471
/gene="NMB0658"
5211. 5471
/gene="NMB0658"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF62315.1"
/db_xref="GI:7413437"
/translation="MLPRLHGLGNGRRLKSKRHPSCRLRMSTGPOPCRRNRRLQ
HRQSRKYGKMKTEIKADGRKVVNRIDSTBADLLPYAGQ"
5502. 5903
/gene="NMB0659"
5502. 5903
/gene="NMB0659"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
Query Match 14.2%; Score 237.2; DB 1; Length 10057;
Best Local Similarity 59.7%; Pred. No. 9.7e-51;
Matches 417; Conservative 0; Mismatches 278; Indels 3; Gaps 1;
QY 35 ccatactgcaagtgtgctgctgcatgcatgcaacgctcagatttgcaaacgattctt 94
DB 2343 CCGTAGCGGGCGGCGGCACATCATACAGCCGCCCTCGGGGGGACTTGGCGCAAGACCCGT 2402
QY 95 ttatccgcaagttctcgacgctcaacatttgaaacgcgaagaaataccactattcg 154
DB 2403 TCATTACCGATTAACGCCCAAGCGCACACTACGACCCGGCGCAATATACACCTCTTCG 2462
QY 155 gcaagcagggggaac---ttgcccagcgcaagcgatcatcagatttggaataatacaaa 211
DB 2463 GGGACCGGGCGGCGGACAGCTTCCGACCGCACCGGCAAAATCAACGTATCAAGACTATA 2552
QY 212 gccatcagttggcaacctgttcattccagcaagcgcgccattaaaggaataatcgctaca 271
DB 2523 CCCACCAATGGGCAACCTGCTCATCCACAGGCAAAATCAACGGGCAACATTCGGGTACC 2582
QY 272 ttgtccgcttccgatacgcaagcgcaagtcattccctccctcgacaacatcgctcac 331
DB 2583 ACACCCGCTTTCCGGACACGACACGAAAGAACACGCCCTTCGACAAACACGCGCGCG 2642
QY 332 attccgattctgatgaagcgcgtagtcccgcttgacgattcagccttiacgcattccat 391
DB 2643 ACAGCGCGAGCGCAAGAAAGCAACGTTGACGAAGGCTTTACCGTATACCGGCTCACT 2702
QY 392 gggacgagtaagaacacatcccgccgacgagctatgaagggccacagcgcgctatc 451
DB 2703 GGGAAAGGACACGACATATCCCGCGATGCTACGACGGCGCGGAAGGCGGCAATTATG 2762
QY 452 ccgctcccaagcgcgagggatataatcagctagacataaaggcgttgcctccaaata 511
DB 2763 CCAAACTTACGGGCGACAGACGATACCTATCATCGTCAACGCGACAGCCCGCAGTA 2822
QY 512 tcgcactaacctgacgacaaacgcgacgacgacgacgacgacgacgacgacgacgacgac 571
DB 2823 TCAAACTCAATTCGACGACACCGCGACGATTCGCGCAACGATTCGCAATTACAGTA 2882
QY 572 ataccgtagtatgtctgacgcaagagtagcgacgattcaaacgcgacccgatatca 631

Db 2883 ACCTGGCAGCAATTTTCACCGCGCGGCAACCAACAAATAAATTTGGACACA 2942
Qy 632 gccccagctgacagatcggaatgcccgcgaagcttcaacggaactgcagatcg 691
Db 2943 ATGCCAAGCTCGACGGCTGGGCGGACAGCATGAGTTTATCAACGCGGCGCGCG 3002
Qy 692 tcaaaatcatctcgccgcgcgcgaagaaatgctggc 729
Db 3003 CGCTCAACCCCTTATTCACGCGCGGCGGCAAGCCGTTGAC 3040

RESULT 28
A96138
LOCUS A96138 1497 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 171 from Patent W09924578.
ACCESSION A96138
VERSION A96138.1 GI:6779961
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1497)
AUTHORS Pizsa,M., Scarlato,V., Rappuoli,R., Grandi,G. and Maignani,V.
TITLE Neisseria antigens
JOURNAL Patent: WO 9924578-A 171 20-MAY-1999;
PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);
CHIRON SPA (IT); GRANDI GUIDO (IT); MALIGNANI VEGA (IT)
FEATURES
source 1. 1497
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 475 a 319 c 365 g 338 t
ORIGIN

Query Match 7.3%; Score 121.8; DB 6; Length 1497;
Best Local Similarity 66.7%; Pred. No. 1e-20;
Matches 174; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 237 ccagcagcgccgcatcaaaagaataatcgctacatgctccgcttccgatacagcgca 296
Db 243 CGAACGAGACGAGATTGAAAGGTATCGGCTATGAACCCATTTTTCAGACACGAGACA 302
Qy 297 cgaagtcacatcccccttcgacaaacatgctcaccatctcgatctgtagagccgtag 356
Db 303 CGAAGTACACAGTCCGTTGATGATTCATGATTCATAAAGCACTTCTGATTCAGCGCGG 362
Qy 357 tccgcttgacggaattcagccttaccgcatcattggaacgatacgaacacacccgcg 416
Db 363 CGTAGACGGCGGCTTTTACCGTTTACCACTTCATCGACAGGGTGGAAATATACATCCCGC 422
Qy 417 cgaagctatgacgagcgacagcgagcgatcctccgctcccaaaagcgcgagagatat 476
Db 423 AGACGATATGACGGCGCTCAAGGCGGCTATTCGGAACCAAGGCGCAAGGATAT 482
Qy 477 atacagctacacataaagg 497
Db 483 ATACAGCTACCATATCAAAAG 503

RESULT 29
A61829
LOCUS A61829 1744 bp DNA linear PAT 09-MAR-1998
DEFINITION Sequence 9 from Patent W09711181.
ACCESSION A61829
VERSION A61829.1 GI:3715998
KEYWORDS
SOURCE Neisseria gonorrhoeae.
ORGANISM Neisseria gonorrhoeae
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 1744)

AUTHORS Meyer,T.F., Rudel,T., Scheuerpflug,I., Fischer, Eckhard and
Maier,J.
TITLE NUCLEIC ACID MOLECULES WHICH CODE PROTEINS WHICH MEDIATE THE
JOURNAL ADHESION OF NEISSERIA CELLS TO HUMAN CELLS
PATENT: WO 9711181-A 9 27-MAR-1997;
COMMENT MAX PLANCK GESELLSCHAFT (DE)
OTHER PUBLICATION DE 19534579 970320.
FEATURES
source 1. 1744
/organism="Neisseria gonorrhoeae"
/strain="MS11"
/db_xref="taxon:485"
/clone_11b="GENOMISCHE BIBLIOTHEK IN PBA"
/clone_11b="(B) CLON(E): H1967/PES25"

BASE COUNT 560 a 383 c 405 g 396 t
ORIGIN

Query Match 7.3%; Score 121.8; DB 6; Length 1744;
Best Local Similarity 66.7%; Pred. No. 1e-20;
Matches 174; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 237 ccagcagcgccgcatcaaaagaataatcgctacatgctccgcttccgatacagcgca 296
Db 284 CGAACGAGACGAGGTTGAAGCATTTATCGGTTATGAACCCATTTTCAGACACGAGACA 343
Qy 297 cgaagtcacatcccccttcgacaaacatgctcaccatctcgatctgtagaagccgtag 356
Db 344 CGAAGTACACAGTCCGTTGATGATTCATGATTCATAAAGCACTTCTGATTCAGCGCGG 403
Qy 357 tccgcttgacggaattcagccttaccgcatcattggaacgatacgaacacacccgcg 416
Db 404 CGTAGACGGCGGCTTTACCGTTTACCACTTCATTCGACAGGGTGGAAATATACATCCCGC 463
Qy 417 cgaagctatgacgagcgacagcgcgatcctccgctcccaaaagcgcgagagatat 476
Db 464 AGACGATATGACGGCGCTCAAGGCGGCTATTCGGAACCAAGGCGCAAGGATAT 523
Qy 477 atacagctacacataaagg 497
Db 524 ATACAGCTACCATATCAAAAG 544

RESULT 30
A61821
LOCUS A61821 3287 bp DNA linear PAT 09-MAR-1998
DEFINITION Sequence 1 from Patent W09711181.
ACCESSION A61821
VERSION A61821.1 GI:3715995
KEYWORDS
SOURCE Neisseria gonorrhoeae.
ORGANISM Neisseria gonorrhoeae
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 3287)
AUTHORS Meyer,T.F., Rudel,T., Scheuerpflug,I., Fischer, Eckhard and
Maier,J.
TITLE NUCLEIC ACID MOLECULES WHICH CODE PROTEINS WHICH MEDIATE THE
JOURNAL ADHESION OF NEISSERIA CELLS TO HUMAN CELLS
PATENT: WO 9711181-A 1 27-MAR-1997;
COMMENT MAX PLANCK GESELLSCHAFT (DE)
OTHER PUBLICATION DE 19534579 970320.
FEATURES
source 1. 3287
/organism="Neisseria gonorrhoeae"
/strain="MS11"
/db_xref="taxon:485"
/clone_11b="GENOMISCHE BIBLIOTHEK IN PBA"
/clone_11b="(B) CLON(E): H1967/PES25"

BASE COUNT 1017 a 740 c 752 g 778 t
ORIGIN

Query Match 7.3%; Score 121.8; DB 6; Length 3287;
Best Local Similarity 66.7%; Pred. No. 1.2e-20;

Matches	174: Conservative	0: Mismatches	87: Indels	0: Gaps	0:
QY 237	ccgaagcgcgcgcacctaataagaataatcgcgttaactctgcgctttctccgatcaaggca	296			
Db 1827	CGAACGGACAGCGGTTTGAAAGGCAATTATGCGTTATGTAAGAACCCATTITTCAGACACGGACA	1886			
QY 297	cgaagtcacatcccccctcgcaacacacatgcctcaattccgattctgaatgaagccgtag	356			
Db 1887	CGAAGTACACAGATCCGCTGTGTAATCAATGATTCAAAGAAAGCACTTGTGATTCAGCGCGG	1946			
QY 357	tcoccttgaagattcagccttaccgcatctccatctgaggaagatacgaaccaccatccgc	416			
Db 1947	CGTAGACGCGCGGTTTACCGTTTACCACTTCATCGACAGSGGTGCGAAATACATCCCGC	2006			
QY 417	cgacgcctatgacgcgcacacagcgcgcgctatccgcctcccaaggcgcgagagat	476			
Db 2007	AGACGGAATATACGGGCGCTCAAGGCGCGGCTGTTATCCGACACACCAAGGCGCAAGGATAT	2066			
QY 477	atacagctacgacataaag	497			
Db 2067	ATACAGCTACCAATCAAGC	2087			
RESULT 31	AF142582	3394 bp	DNA	linear	BC1 18-MAY-1999
LOCUS	AF142582				
DEFINITION	Neisseria gonorrhoeae maf gene cluster, complete sequence.				
ACCESSION	AF142582				
VERSION	AF142582.1	GI:4836552			
KEYWORDS					
SOURCE	Neisseria gonorrhoeae.				
ORGANISM	Neisseria gonorrhoeae.				
REFERENCE	1 (bases 1 to 3294)				
AUTHORS	Eickenjaeger,S., Meyer,T.F., Fischer,E., Maier,J., Manning,P.A.,				
TITLE	Rudel,T., Scheuenpflug,I., Schulz,E. and Schwan,E.T.				
JOURNAL	Submitted (13-APR-1999) Molekulare Biologie, MPI fuer				
FEATURES	Intektionsbiologie, Monbijoustr. 2, Berlin 10117, Germany				
source	Location/Qualifiers				
gene	1..3294				
CDS	/organism="Neisseria gonorrhoeae"				
	/strain="MS11"				
	/db_xref="taxon:485"				
	143..457				
	/gene="mafI"				
	143..457				
	/gene="mafI"				
	/codon_start=1				
	/transl_table=1				
	/product="MafI"				
	/protein_id="AAD31037.1"				
	/db_xref="GI:4836553"				
	/translation="MRSRSLSRYSRQNKLIIEFVAGVARTAAEPDSIYYTDCRYRDY				
	RF"				
	LDAEFGFHRFNHSTHFAEGRNHNLINGIFENFRAKRIHUKRDGIKHEFELYLNECE				
gene	590..1552				
CDS	/gene="mafA"				
	590..1552				
	/gene="mafA"				
	/codon_start=1				
	/transl_table=1				
	/product="Adhesin MafA"				
	/protein_id="AAD31038.1"				
	/db_xref="GI:4836554"				
	/translation="MRAKLIPILFSVFITISACGTLTGIPSHGGAKREAVQELVAAS				
	ARAAVKMDLDALHGKVALYIATMGDSGSLTGGRTSIDALRGETINSVAVTDVAAS				
	TIPTPEATFTSGGLTGILTSLTNAPALSROSDSGSRSLGILNIGMGYRYNNA				
	TLTNPRLATLSLHVQVPEFLRGIDIVSPANAPTDYVINTIDVGTINRPMHLYNN				
	ELTKAOKIEYFAVDRTKKLLIKPKINAFAPAAAKENALTMGPYKSKGIKPIEGLM				
	VDFSDIRPYGNHTGNSAPSVADNSHSEYTGISDAVYRHHROGP"				
	1592..3121				

Query Match	Best Local Similarity	66.7%	Score 121.8	DB: 1	Length 3294
Matches 174	Conservative	0	Mismatches 87	Indels	Gaps 0
QY 237	ccagcagcgccgcatlaagaaatagcgtacatgctccgcttccgcgtacacgagca 296				
Db 1834	CGAACGCGACAGGGTTTGAAGCATATATGCGTTATGAAACCCATTWTCAGACACGACGA 1893				
QY 237	cgaagtcattcccccctcgagaaacattgcctcacatccgattctgtagagcgag 356				
Db 1894	CGAAGTACACAGCTCGTTCGTAATCAATGATTCACAAAAGCACTTCGTGATTCAGGCGCG 1953				
QY 357	tcocgctgacgattcagcgtcttaccgcatccatcgtatggagacgatacagacacatccgcg 416				
Db 1954	CGTAGACGCGCGATTTCACGTTTACCACTTCATCGACAGGATCGGAAATATACATCCCGC 2013				
QY 417	cgagcgcatatagcgagcagagggcgagcgtatccgctccgcaaaagcgagagatg 476				
Db 2014	AGACGGATATAGCGGGCTCTCAAGGGCGGCTTATCCGAGACACCAAGGCGCAAGGATAT 2073				
QY 477	atcacgctacgacataaag 497				
Db 2074	ATACAGCTACCATATCAAGAAG 2094				
RESULT 32	NMA122491/c	340806 bp	DNA	linear	BCT 04-DEC-2000
LOCUS	NMA122491				
DEFINITION	Neisseria meningitidis serogroup A strain Z2491 complete genome; segment 1/7.				
ACCESSION	AL162752	AL162752			
VERSION	ALI62752.2	GI:73787878			
KEYWORDS	Neisseria meningitidis Z2491.				
SOURCE	Neisseria meningitidis Z2491.				
ORGANISM	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.				
REFERENCE	1 (bases 1 to 340806)				
AUTHORS	Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C., Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T., Davies,R.M., Davis,P., Devlin,K., Feltham,T., Hamlin,N., Holtgrave,S., Jagers,K., Leather,S., Moule,S., Mungall,K., Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M., Skellon,T., Whitehead,S., Spratt,B.G. and Barrall,B.G.				
TITLE	Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491				
JOURNAL	Nature 404 (6777), 502-506 (2000)				
MEDLINE	20022556				
REFERENCE	2 (bases 1 to 340806)				
AUTHORS	Parkhill,J.				
TITLE	Direct Submission				

JOURNAL Submitted (30-MAR-2000) Submitted on behalf of the Neisseria sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk

COMMENT Notes: Details of N. meningitidis sequencing at the Sanger Centre are available on the World Wide Web. (URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).

FEATURES

source location/Qualifiers

1. 340806

/organism="Neisseria meningitidis 22491"

/strain="22491"

/db_xref="taxon:122587"

/note="serogroup: A"

/complement(38. .358)

/gene="NMA0001"

/complement(38. .358)

/gene="NMA0001"

/note="NMA0001, unknown, len: 106 aa"

/codon_start=1

/transl_table=11

/product="hypothetical protein NMA0001"

/protein_id="CAB83321.1"

/db_xref="GI:7378779"

/db_xref="SPTREMBL:Q9JX93"

/translation="MSNTQIRSADEFTTSRWAGALDIANNNGTTRVRLHMTDQPYKWHVNDGEVFAVMGEVDHMYRENGEHHVRLKSGDIFVAGIGTEHVAHPRGRTRLVTEK EGSV"

/complement(414. .2438)

/gene="nuoL"

/complement(414. .2438)

/gene="nuoL"

/EC_number="1.6.5.3"

/note="NMA0002, nuoL, NADH dehydrogenase I chain L, len: 674 aa; similar to many e.g. NUOL_RHOCA NADH dehydrogenase I chain L (EC 1.6.5.3) (700 aa). Contains Pfam match to entry PF00361 oxidored_q1, NADH-ubiquinone/plastoquinone (complex I), various chains, and to entry PF00662 oxidored_q1.N, NADH-ubiquinone oxidoreductase (complex I), chain 5 N-terminus"

/codon_start=1

/transl_table=11

/product="NADH dehydrogenase I chain L"

/protein_id="CAB83322.1"

/db_xref="GI:7378780"

/db_xref="SWISS-PROT:Q9JX92"

/translation="MNDWTLXLIILALVPLAGSLIAGFGNKGIRAGAHVYTIIGVAVS AVLSAVYLMGFLNGSRKAFDENYTVLIMGGLDPSVGFIVDTMTAMMAYVYGSIMV HVTYIGVMDEKVGQRFPSYISLFTFSMLMLIMNNFIQLFGEVAVGLVSYLIGF YKRRSATFANLKAFLINRVGDFGLGILVLAFGSLRYQDVFAVLYPNQVATIQ LFGVEMSLITVTCLLFLVGAMGSAQFPLHVMIPDSMEGPTPISALIHAAVTAGL FVSRMSPTIEMSTALSYIMVIGATILEMGLVIONDIKRVAVSTLSQLEMTV ALGASASVAMRHVMTHAFKALFLFAAGSALIGMHHDDMRHNGKKMPTWLM LIGNSLITGTPFSGFYSKDSITTEAKYSTLPSCGRYFVAVLSVPTATYAFQYFM VHGEEKMRSLEPHSHSDGHEHGHGKNDNPRESPLVTLPLTLAVPSVITGIAL EPMLYGDFEKDVIENADAPHTMIHKEEFHGLAMVSHLSHSPVILALAGVLSAML LVKLPHLPAKIAQAFRPVYVLEENKYYDALYENFAKCTRALGTFEFMYGDAIID NCIIVGSAVLGVAAQVARKVYGTCTYTAAMAVEGVLVLMGTWGLER"

1035. 1044

/note="Core DNA uptake sequence: gccgtctgaa"

/label=DUS

/complement(1116. .2030)

/gene="nuoL"

/note="Pfam match to entry PF00361 oxidored_q1, NADH-ubiquinone/plastoquinone (complex I), various chains, score 351.80, E-value 7.3e-102"

/complement(2061. .2243)

/gene="nuoL"

/note="Pfam match to entry PF00662 oxidored_q1.N, NADH-ubiquinone oxidoreductase (complex I), chain 5 N-terminus, score 77.20, E-value 3.5e-19"

/complement(2487. .2828)

/gene="NMA0003"

/complement(2487. .2828)

CDS

/gene="NMA0003"

/note="NMA0003, unknown, len: 113 aa"

/codon_start=1

/transl_table=11

/product="hypothetical protein NMA0003"

/protein_id="CAB83323.1"

/db_xref="GI:7378781"

/db_xref="SPTREMBL:Q9JX91"

/translation="MMDRKNOLEOFEFHKKMLNIYORALNIPQYKATRFQIVNEFGK EADKLSTGEKKTQGTFTELLSCGCVHAKYSMEYLVLRKPMQDLTFEOLAVARK RLRFVGFVSPK"

/complement(2833. .2837)

/complement(2932. .3527)

/gene="NMA0004"

/complement(2952. .3527)

/gene="NMA0004"

/note="NMA0004, unknown, len: 191 aa; similar to hypothetical proteins e.g. Y977_HA9N HI0977 (191 aa), fasta scores; E(): 0; 80.5% identity in 190 aa overlap, and to proteins involved in cell cycle e.g. TR:054679 (EMBL:AF036487) Lactococcus lactis Plasmid pM24000 (EMBL:AF036487) Lactococcus lactis Plasmid pM24000 putative mobilization protein (200 aa), fasta scores; E(): 0, 55.1% identity in 167 aa overlap, and FIC_ECOLI cell filamentation protein FIC (200 aa), fasta scores; E(): 0.093, 29.4% identity in 126 aa overlap"

/codon_start=1

/transl_table=11

/product="hypothetical protein NMA0004"

/protein_id="CAB83324.1"

/db_xref="GI:7378782"

/db_xref="SPTREMBL:Q9JQ89"

/translation="MPSNPICKTKMSIDEOGSLHNRRLFESGDDIRIEVGTAGLQO IHRVFLGLYFAGQIREDNISKGFREFANAMYLKALVKIQMPERTEETIAKVE MNIAHPLFEGNGRSTRIMDLVLRKRLKRWVWQVNSKTYLQAMERSVNDLEFL LKNDITDVNDNEILFEKIGDSYVYEGYKKG"

/complement(3516. .3523)

/gene="NMA0004"

/note="Core DNA uptake sequence: gccgtctgaa"

/label=DUS

/complement(3561. .3866)

/gene="nuoK"

/complement(3561. .3866)

/gene="nuoK"

/EC_number="1.6.5.3"

/note="NMA0005, nuoK, NADH dehydrogenase I chain K, len: 101 aa; similar to many e.g. NUOK_RHOCA NADH dehydrogenase I chain K (EC 1.6.5.3) (102 aa), fasta scores; E(): 1.1e-24; 65.0% identity in 100 aa overlap. Contains Pfam match to entry PF00420 oxidored_q2, NADH-ubiquinone/plastoquinone oxidoreductase chain 4L"

/codon_start=1

/transl_table=11

/product="NADH dehydrogenase I chain K"

/protein_id="CAB83325.1"

/db_xref="GI:7378783"

/db_xref="SPTREMBL:Q9JQ09"

/translation="MITTLNHYLVGLALFGISAMGIFMNRKNVLVILMSIELMLAVN FNFIAFSOHLGDTAGQIEFVFYLVAAASALGIALIVLYRRROTINVDLELKG"

/complement(3564. .3863)

/gene="nuoK"

/note="Pfam match to entry PF00420 oxidored_q2, NADH-ubiquinone/plastoquinone oxidoreductase chain 4L, score 141.50, E-value 1.5e-38"

/complement(3863. .4534)

/gene="nuoJ"

/complement(3863. .4534)

/gene="nuoJ"

/EC_number="1.6.5.3"

/note="NMA0006, nuoJ, NADH dehydrogenase I chain J, len: 223 aa; similar to many e.g. NUOJ_RHOCA NADH dehydrogenase I chain J (EC 1.6.5.3) (202 aa), fasta scores; E(): 9.7e-21; 41.3% identity in 201 aa overlap. Contains Pfam match to entry PF00499 oxidored_q3,

CDS

		NADH ubiquinone/plastoquinone oxidoreductase chain 6"
		/codon_start=1
		/product="NADH dehydrogenase I chain J"
		/protein_id="CAB83326.1"
		/db_xref="GI:7378784"
		/db_xref="SPTREMBL:Q9JX90"
		/translation="MTFQLILFYIFAVYIILGYAKITVYAKNPVAALHLVLTPCVSAMLMIMQAEEFLGTVLVVVYGAVMYLFVVMALMIDIEEMAGFRMRHPVAVGGTLLAAVALILINRKTDLAAGFKMDIPADNNTRDGSRTYTLLPFELAAVILLGWAAIALIVHRKTVNPKRMDDPDQVKVRADQGRRRLVTKMEAVKRQTSAESESVDLKKKEEGA"
RBS		complement(3874..3879)
		/gene="nuoJ"
misc-feature		complement(4034..4525)
		/gene="nuoJ"
		/note="Pfam match to entry PF00499 oxidored_q3, NADH-ubiquinone/plastoquinone oxidoreductase chain 6, score 77.90, E-value 2e-19"
gene		complement(4570..5277)
		/gene="NMA0007"
CDS		complement(4570..5277)
		/gene="NMA0007"
		/note="NMA0007, unknown, len: 235 aa"
		/codon_start=1
		/transl_table=1
		/product="hypothetical protein NMA0007"
		/protein_id="CAB83327.1"
		/db_xref="GI:7378785"
		/db_xref="SPTREMBL:Q9JX89"
Query Match	6.6%; Score 110.2; DB 1;	Length 340806;
Best Local Similarity	62.5%; Freq. No. 3.5e-17;	Mismatches 103; Indels 0; Gaps 0;
Matches 172; Conservative 1;	Mismatches 103;	Indels 0; Gaps 0;
Oy	237	ccagcaggcggcataagaagaatcatgctcacattgccgcttttcgcgtacacggcca 296
Db	305687	CGAACGGACAGGATTGGAAGGCATTTATGAAACCATTCTTCAGCAAGGACA 305628
Oy	297	cgaatccattcccccttgcgaaccatgacctcaatcttcgatagtgaagccggttag 356
Db	305627	TGAAGTACACACTCCGTTCGATAATCATGATTCAAAAGCACTTGATTCAGCGGCG 305568
Oy	357	tcccgttgacgattcaagcctttacgcgatccatctggagcagatcagacaacatccgcg 416
Db	305567	CGTAGACGGTGGTTTACCGGTTTACCACTTCATCGACAGGCGTGGAAATCATCCGGA 305508
Oy	417	cgagcgctatgacggcgccacggcgggcggcctatccgcgtctcccagaagcgagggatat 476
Db	305507	GGATGGATATGACGGGGCGCAAGGACGAGTAATTCGCCGCCCGGAGGACGAAGGATAT 305448
Oy	477	atacagctacgacataaaggcgltgcaccaata 511
Db	305447	ATACAGCTACTATGTCCAAGGAACCTTCAACAAAA 305413
RESULT 33		
LOCUS	A96134	1449 bp DNA linear PAT 07-SEP-2000
DEFINITION	Sequence 167 from Patent WO924578.	
ACCESSION	A96134	
VERSION	A96134.1	GI:6779959
KEYWORDS	.	
SOURCE	unidentified.	
ORGANISM	unclassified.	
REFERENCE	1 (bases 1 to 1449)	
AUTHORS	Piazza,M., Scariato,V., Rappuoli,R., Grandi,G. and Masiagnani,V.	
TITLE	Neisserial antigens	
JOURNAL	Pizza MARIAGAZZA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT); CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)	
FEATURES	Location/Qualifiers	

[illegible]

|||||
Db 423 GGATGGATATGACGGCGCCGACGAGCATATTCGCCGCCCGGAGGACGACGATAT 482
Oy 477 atacgactacgacataaaggcgtgcccaata 511
Db 483 ATACAGCTATATGTCACAGAACTTCACACAAAA 517

RESULT 35
AE002559
LOCUS 10869 bp DNA linear BCT 25-MAY-2000
DEFINITION Neisseria meningitidis serogroup B strain MC58 section 201 of 206
of the complete genome.
AE002559 AE002098
VERSION AE002559.2 GI:7413482
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
COMMENT
FEATURES
source

1 (bases 1 to 10869)
Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
Dodson,R.J., Nelson,W.C., Gwinn,M.L., Deboy,R., Peterson,J.D.,
Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Maignani,V.,
Piazza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
Rappuoli,R. and Venter,J.C.
Complete genome sequence of Neisseria meningitidis serogroup B
strain MC58
Science 287 (5459), 1809-1815 (2000)
10710307
2 (bases 1 to 10869)
Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
Dodson,R.J., Nelson,W.C., Gwinn,M.L., Deboy,R., Peterson,J.D.,
Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Maignani,V.,
Piazza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
Rappuoli,R. and Venter,J.C.
Direct Submission
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
On Apr 4, 2000 this sequence version replaced gi:7227357.
Location/Qualifiers
1..10869
/organism="Neisseria meningitidis MC58"
/strain="MC58"
/db_xref="taxon:122586"
/note="serogroup: B"
150..524
/gene="NMB2095"
150..524
/gene="NMB2095"
/note="similar to PID:1817528 percent identity: 53.99;
identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="adhesin complex protein, putative"
/protein_id="AAF42412.1"
/db_xref="GI:7227358"
/translation="MKLLTVAIISSAIALSSMAAACTDNEPTVAKKTVSVYCCGQKKY
KTYVGNKQGLITTYASAVKRWMPVNLDSNVETFFYEGGYVLTGYVMGKSY
RKQPIMITADNQIVFKDCSPR"
complement(712..2178)
/gene="NMB2096"
complement(712..2178)
/gene="NMB2096"

gene
CDS

/note="similar to SP:P33940 PID:424051 GB:U00096 percent
identity: 69.73; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="male:quinone oxidoreductase"
/protein_id="AAF42413.1"
/db_xref="GI:7227359"
/translation="MAEANDVVLVGGGIMVSATLGVTLKELPSWETLLIERDALE
SSNMNNAAGTSALCELVAPLGNGITDPARALNIAOPHYRQFAPATVAEGLK
DNEFTNAVPHMSLVNNEHCSTLQRYDAFKYQKLEENEFSTDRKISDMAPLMAR
RDEQNPVANSABEGTDVDEGLTRQMYVYLGQKGVKTEFNHVEDIKRESDAWLK
TADRENPDGOLITRTFLFLGAGGALTLLQSGIPEGAGYGFVPSGLFERNBET
AEQNAKYVGOASVAPRPSVYHLDRNVGKRHLMEFGYAGFSNFLQGSIMLPL
SIHMNLVPLMCAGANMPLTKYTLGELKTEDEERPSALILEYYPEANPDQMLITAGO
RVQITKDSKSGVYQFGTEIYAHADGSAIALILGASPGASTAVPLMIRLMHQCPEPRA
PSWEDRLKELVPGIGIKLENPERADEITATYAKVDT"
complement(2230..2319)
/gene="NMB2097"
complement(2230..2319)
/gene="NMB2097"
complement(2230..2319)
/gene="NMB2097"
complement(2230..2319)
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42414.1"
/db_xref="GI:7227360"
/translation="MLATHSNPETAISDTHHCRFSLLAFRL"
complement(2715..3305)
/gene="NMB2098"
complement(2715..3305)
/note="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF42415.1"
/db_xref="GI:7227361"
/translation="MSEKIIIVLSIGVRLPVDVHNETGLREAVCDGKWLKGYTS
AHPHRAVATYCTPALATRIAPFVAVDEBACRVGTGTYHFEFPIPLDIDGFTYMASA
RTRINTCKCKMLLDSAPFVLAICRCGWRDITLNLASQRIERLGAEKDGVLMHMLR
KGTGYVDIVVYMLREDCKNREILITGLAGVGY"
complement(3366..3947)
/gene="NMB2099"
complement(3366..3947)
/gene="NMB2099"
complement(3366..3947)
/note="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF42416.1"
/db_xref="GI:7227362"
/translation="MRNEKRRLRRELGRSSOMGDRVRAATVKNHLLKRYTKKR
KIGYVWPMGKELRLDGFVRAAQGRGAELVLYEPRSRRMFTPPYDGVQEKRRR
AKLHVQFAGRRKRVYDLNLLPVVGMPLGYRLQAGGYDATLSAMKYRLQAKTV
GVGFACOLVRLPVEAHDRSLDGFVSEAGILCP"
4025..4258
/gene="NMB2100"
4025..4258
/gene="NMB2100"
4025..4258
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF42417.1"
/db_xref="GI:7227363"
/translation="MWHIVAGLYFAVMVSAOPSTARALIVLFWAVLPFTVFTPT
ITVRRNHLMRQEOEAESDQRRQROKDSGTRK"
4431..5159
gene
CDS

gene
CDS
/gene="NM2101"
4431..5159
/note="similar to SP:P34831 GB:X53651 P1D:296030 percent identity: 78.02; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="30S ribosomal protein S2"
/protein_id="AAF42418.1"
/db_xref="GI:7227364"
/translation="MSQITRMQIEGVHGHGTRRWNPMAQYIGARKIHVILE KILPMQDAQEVRLVANKGYLVETGRQARDI;REEATRAGMPEVDYRWLGMLT NKTKVQSIKRLKEETKALENAESEGSKLEIMQDVKEKSLGKINMGGLPDA IFEVDYQKGTLEVEAKELGIVIAVVDNTNSPDGKYVIPGNDSSAKAIRLCRGIA DAVLEKKNALQETVAAQEAALAE"
5288..6142
/gene="NM2102"
5288..6142
/note="similar to GB:D26562 SP:P02997 GB:D13334 GB:V00343 P1D:1208943 percent identity: 69.01; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="elongation factor TS (EF-TS)"
/protein_id="AAF42419.1"
/db_xref="GI:7227365"
/translation="MAEITAKMVADLRATGIGMCKKALVAEENFKAETLRK SGAKGKLAGRTAAGSVLAIATNGVAGALVERNCETDEYAKDGEFEAFVAKTAE KKPASVEISELVEARKEATIKLGNMSVRRPOVDTNOQVAVIHGALATVAVK VIKGSDVARKIGIMHVAAPKPCVSEAEVDAVEKRRHYETEDALISGKPADIAAKV EGRIRKFLAETILNGCAVFMNDQTVAFQSKENGTEVISEFRYKVDGIEKKAVDYAA EVAAAKAY"
6359..7078
/gene="NM2103"
6359..7078
/note="similar to GB:D26562 SP:P29464 GB:D13334 GB:X78809 P1D:1208944 percent identity: 73.16; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="uridylylate kinase"
/protein_id="AAF42420.1"
/db_xref="GI:7227366"
/translation="MTQOIKYKRVLLKISGESLMSDPRGINHDTIVQTVGEIAEVK MGVOGIVGGGNIERGVSAQASMDRATADMGKATMNLALKADEFETGIRARV OSALSMOIAETIYARPKAIQYIEEGKVYIFAGTGPEFTTIDPAALRGAEKMDVWL KATINVDGYTADPKDPSATRIETITFDALILKMKVMDATAFALCRERKLNIVFGI AKESGLKRVITGDEEGTLVHC"
7262..8225
/gene="NM2104"
/note="similar to GP:4838554 percent identity: 99.38; identified by sequence similarity; putative"
8265..9716
/gene="NM2105"
8265..9716
/note="similar to GP:4838555 percent identity: 0.00; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="mafB protein"
Query Match 6.4%; Score 107; DB 1; Length 10869; Best Local Similarity 61.8%; Pred. No. 1.1e-16; Matches 170; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
QY 237 ccagcagcgccgcaataaagaatatacgtcgtcgtttccgcatcagcgaca 296
DB 8507 CGAACGACAGGGTTTGAAGGTTATCGTTATGAACCCATTTTTCAGGGCAGGACA 8566

QY 297 cgaagtlcaatcccccctcgacaacatgcctacatccgatctgataagccgtag 356
DB 8567 TGAAGTACACAGTCCGTTGCATCATATGATTCAAAAAGCACTTCTGATTCAGCGCG 8626
QY 357 tcccggttcagcgattcagccttcacgcacatccatcttgagcagtaacacatccgc 416
DB 8627 TGTAGACGCGCGTTTACTGTTTACCACTTCATCGAACAGGGGTGCGAAATTCATCCGGA 8686
QY 417 cgaagcgtatagcgagcgacagcgcgcgatcccgctcccaaggcgagggatgata 476
DB 8687 GGATGGATATGACGGGCGCGAAGCAGCATTTATCCGCCCGGAGGACCAAGGATAT 8746
QY 477 atacagctacgacataaagcgcttgcccaata 511
DB 8747 ATACAGCTATTATGTCACAAAGCACTTCACAAAAA 8781
RESULT 36
AX044035 172325 bp DNA linear PAT 24-NOV-2000
LOCUS Sequence 114 from Patent WO0066791.
DEFINITION AX044035
ACCESSION AX044035
VERSION AX044035.1 GI:11342919
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
REFERENCE
AUTHORS
1 (bases 1 to 172325)
Plaza,M., Hickey,E., Peterson,J., Tettelin,H., Venier,J.C., Maignan,I.V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M., Scarlato,V., Rappuoli,R., Frazer,C.M. and Grandi,G.
TITLE
JOURNML
Neisseria genomic sequences and methods of their use
Patent: WO 0066791-A 114 09-NOV-2000.
FEATURES
location/Qualifiers
source
1..172325
/organism="Neisseria meningitidis"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces. -seq 1: 1 to 349980 349980 bases-seq 108: 300001 to 649980 349980 bases-seq 109: 600001 to 949980 349980 bases-seq 110: 900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980 349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq 113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to 2272325 172325 bases"
BASE COUNT 43072 a 47583 c 41465 g 40205 t
ORIGIN
Query Match 6.4%; Score 107; DB 6; Length 172325; Best Local Similarity 61.8%; Pred. No. 2.1e-16; Matches 170; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
QY 237 ccagcagcgccgcaataaagaatatacgtcgtcgtttccgcatcagcgaca 296
DB 122937 CGAACGACAGGGTTTGAAGGTTATCGTTATGAACCCATTTTTCAGGCGACGACA 122996
QY 297 cgaagtlcaatcccccctcgacaacatgcctacatccgatctgataagccgtag 356
DB 122997 TGAAGTACACAGTCCGTTGATCATGATTCAAAAAGCACTTCTGATTCAGGCGCG 123056
QY 357 tcccggttcagcgattcagccttcacgcacatccatccatcgaagcgatagcagaccatccgc 416
DB 123057 TGTAGACGCGCGTTTACTGTTTACCACTTCATCGAACAGGGGTGCGAAATTCATCCGGA 123116
QY 417 cgaagcgtatagcgagcgacagcgcgcgatcccgctcccaaggcgagggatgata 476
DB 123117 GGATGGATATGACGGGCGCGAAGCAGCATTTATCCGCCCGGAGGACCAAGGATAT 123176
QY 477 atacagctacgacataaagcgcttgcccaata 511
DB 123177 ATACAGCTATTATGTCACAAAGCACTTCACAAAAA 123211

RESULT	37			375 bp	DNA	linear	PAT 07-SEP-2000
LOCUS	A96130						
DEFINITION	Sequence 163 from Patent WO924578.						
ACCESSION	A96130						
VERSION	A96130.1	GI:6779957					
KEYWORDS							
SOURCE	unidentified.						
ORGANISM	unclassified.						
REFERENCE	1 (bases 1 to 375)						
AUTHORS	Pizza,M., Scariato,V., Rappuoli,R., Grandi,G. and Maignani,V.						
TITLE	Neisserial antigens						
JOURNAL	Patent: WO 9924578-A 163 20-MAY-1999;						
	PIZZA MARIAGAZZA (IT); SCARIATO VINCENZO (IT); RAPPUOLI RINO (IT);						
	CHIRON SPA (IT); GRANDI GUIDO (IT); MALIGNANI VEGA (IT)						
FEATURES	Location/Qualifiers						
source	1..375						
	/organism="unidentified"						
	/db_xref="taxon:32644"						
BASE COUNT	105 a 89 c 90 g 90 t 1 others						
ORIGIN							
Query Match	5.9%; Score 98; DB 6; Length 375;						
Best Local Similarity	59.6%; Pred. No. 1.2e-14;						
Matches 164: Conservative	0; Mismatches 11; Indels 0; Gaps 0;						
Qy	237 ccagcagcgcgccattaaagaataatcgcctacattcgcgtttccgatcaaggca 296						
Db	27 CGAAGCGACAGGGTTGAGCGTTATCGGTTATGAACCCATTTTTCAGGGCAGGACA 86						
Qy	297 cgaagtcacatcccttcgcacacacatgcctcaccatccgatcgtatcgtgaagccgtag 356						
Db	87 TGAATGTACACAGTCCGTTGCATCATCATGATTCMAAAGCACTTCTGATTTGAAGGGCGG 146						
Qy	357 tcccgctgacgagatcagccttaccgcatalcatgttggaaggatacgaaccatccgcg 416						
Db	147 TGTACACGGCGGCTTTACTGTTACCTTACCAACTTCATCGAACAATGCTCGAATAATCCATCCGCA 206						
Qy	417 cgaagcgtatgacggggccacagcgcgcgctaccgcgtccgctcaaaaggcgagagatat 476						
Db	207 GGATGAAATGTAGCGGGCCGCAAGCAGCGNATTATATCCGCCCGCGGAGGAGCAAGGATAT 266						
Qy	477 atacagctacgacataaaggcgctgtgcccaata 511						
Db	267 ATACAGCTATTATGTCAAGAAGAACTTCACACAAAA 301						
RESULT 38							
AF429315		125020 bp	DNA	linear	PRI 18-JAN-2002		
LOCUS	AF429315						
DEFINITION	Homo sapiens junctophillin 3 (JPH3) gene, partial cds.						
ACCESSION	AF429315						
VERSION	AF429315.1	GI:17646244					
KEYWORDS							
SOURCE	human.						
ORGANISM	Homo sapiens						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
REFERENCE	1 (bases 1 to 125020)						
AUTHORS	Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,						
	Ingersoll,Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,						
	Potter,N.T., Ross,C.A. and Margolis,R.L.						
TITLE	A repeat expansion in the gene encoding junctophillin-3 is						
	associated with Huntington disease-like 2						
JOURNAL	Nat. Genet. 29 (4), 377-378 (2001)						
MEDLINE	2183737						
PUBMED	11694876						
AUTHORS	2 (bases 1 to 125020)						
	Holmes,S.E., Ingersoll,Ashworth,R.G., Ross,C.A. and Margolis,R.L.						

FEATURES	source	Location/Qualifiers
TITLE	Direct Submission	
JOURNAL	Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA	
repeat_region		1. .125020 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="16" /map="16q24.3, between D16S520 and WI-12410" /note="Isolated from a patient with Huntington's disease-Like 2 (HDL2)" complement(35581. .35746) /rpt_type=tandem /rpt_unit=ctg complement(<36507. .>36887) /gene="JPH3" /product="junctionophillin 3" complement(<36507. .>36887) /gene="JPH3" /note="Jp3" complement(<36507. .36887) /gene="JPH3" /note="component of the junctional complex between plasma membrane and endoplasmic reticulum" /codon_start=1 /product="junctionophillin 3" /protein_id="AA140941.1" /db_xref="GI:17646245" /translation="MSSGGFRNPDDGSGYGGGWEDGKAHGHCTGPKGDEYTSNAG HGFVLYGYTTPSGNTGYGTGAQCKRHIGLESGKAWYKKEHMGFRGVGVECA NGAEYEGWSNGLDDEGTETFTSYSG"
gene		
CDS		
BASE COUNT	29056 a 32731 c 30656 g 28283 t 4254 others	
ORIGIN		
Query Match	3.1%: Score 51.8; DB 9; Length 125020;	
Best Local Similarity	10.6%: Pident. No. 0.053; Mismatches 445; Indels 8; Gaps 3;	
Matches 103; Conservative 414; Predicted 445; Indels 8; Gaps 3;		
262 atcgcgtacatgtcgcgccttcgcgtacgcagcgaagtcacatcccttcgacaaac	321	
Db 16923 MIMSCCMWYYSMMWSSMGCMCMGRRRSKMKWYSRGMSRSKSMRYGSGSKMMRSSMM	16982	
322 catgctcacatcgcgtatctgaagcggtagtccgttgacggatcgagccttac	381	
Db 16983 CTSSCYASMCWCWCSCCCMRSCCCRCRYCCCMRYCCACCKCYMSYTMASYSMSRSY	17042	
382 cgcattcattgagcagatcagacacacaccccgagcagcgtatagcggcacagggc	441	
Db 17043 WKRSKMCWSRMCSSRSKSSKSGCGSGMGKGGKGSYGRKTKRKSRGKMGKAWMYR	17102	
442 ggcgcgtatccgcgtccccaaggcgcgagatataatacagctaacgacataaaggcgtt	501	
Db 17103 RSMRKKMKWYSSKMYCMYCMWRRCYCSYMSRSMCMCSYCYACCKSMCYSCYGGMS	17162	
502 gcccaaatatccgcgtcacaacctgacgcgcaaacgcgagcagcgaacgcgcttgcgc	561	
Db 17163 KGIYACSTRGSSSKYCMRGSTYSISGCGCCCTTTTCCCCANANGGGAAGCTTTMC	17222	
562 cgcttcacaatcccgtagtactgctgacgcaagagtagtagcgagcatcacaagcgcc	621	
Db 17223 NKKTYSYRKRNRGNCAMCKYNNYNNNSWRSSCFRAGMGCCTYKSSMTMSMASYCMCMMYC	17282	
622 acccgatacagcccgagctggaacaatcgggagcaatgcccgcggaagccttcaagcgcat	681	
Db 17283 SMKSMASRGMSWSSYKMKMKSSWRMSCYMKCMCMKMCYCMMSRSGMSSTYMTASMS	17342	
682 ggcagatcgtcaaaaacatcaccgvcgvcgvcgaggaatattcgvcgacagcgatgc	741	
Db 17343 SSMGCTCTRCYMGSSKSCSKSYMMNRSMKMGKMRGMSGMSASASRSCYKCYKCTM	17402	
742 gtcgaggtataagcgaggctcaaacatgtctgtatg--caagcttggtctcgtctt	799	

Db 17403 RCGMMSSKCYRCACGMMKGYRYRMCMGSKRRKRWGKSAAMYRMRKMGKRGAMMCMKC 17462
Qy 800 ccaccgaaacaagatgagcgatcaacagattgtgcagatatgagcgaactcaagaact 859
Db 17463 YSRMSRCMMKMSYSCASCRSCAMMSGTYMCMSSYVCTCMCTGCTSYMYRCTCMCK 17522
Qy 860 atgcgcgagcagccatccgcgcatgtggc---agtcacaaaccccaatgcccgcacaagc 915
Db 17523 GWSYTKRSKSSMSYKGRKRSYMCCTSRGAMSCWRRCCTYMGASMRAGSMRRR 17582
Qy 916 atagaagcgcgtacgaatctcttacgcgcagctatcccccgtcaacagagattgagctgtt 975
Db 17583 AKGRWGGGSKMMWTCGMMRSKYTYCTGRMMATTWCC--CMRRSTMYRSMWAMMR 17640
Qy 976 cggggaataacgagcttgagcgatcagcagcaatcccttcaacgagcgatgcagatggc 1035
Db 17641 KSSWGMRMGSSASRRKSSASRSMCMRMSGSSKMMGGRSRSRASSCKKSR 17700
Qy 1036 gagatcgcatgcccgaagaggaatccgcgcgtcagcagcaatttgcgcgagcagatc 1095
Db 17701 GMRRSKRSKRYRKGRGKRRKRSMTKSGSKGKSCWKRSGSMYSYSSSCMMMS 17760
Qy 1096 gccaaatcccgccctaccatcccgaaatccgtcgaatccatcgatgagcagcgttac 1155
Db 17761 KSCMCCMMKRRACACYSMCTSMYRCMGCMKSYSTGTTCKCTGYKRSRTWMM 17820
Qy 1156 ggcgaagaataacatccctcctcaacgcgtgcgcgcgtcaacggaagaatgtgaactg 1215
Db 17821 YSWTSKMYMRAWAAGCYMSTRMCMMSMCARKGCMCTMGAGATCAAGTAGAGTG 17880
Qy 1216 gcaacacaac 1225
Db 17881 AGATAGAAAC 17890

RESULT 39
AE002514 10947 bp DNA linear BCT 25-MAY-2000
LOCUS Neisseria meningitidis serogroup B strain MC58 section 156 of 206
DEFINITION
ACCESSION AE002514 AE002098
VERSION AE002514.1 GI:7226886
KEYWORDS
SOURCE Neisseria meningitidis MC58.
ORGANISM Neisseria meningitidis MC58.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.

REFERENCE
AUTHORS 1 (bases 1 to 10947)
Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
Dodson,R.J., Nelson,W.C., Gwin,M.L., DeBoy,R., Peterson,J.D.,
Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
Dougenly,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
Cittone,H., Clark,E.B., Cotton,M.D., Uterback,T.R., Khouri,H.,
Oin,H., Yamathayan,J., Gill,J., Scariato,V., Maignani,V.,
Piazza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
Rappoli,R. and Venter,J.C.
Complete genome sequence of Neisseria meningitidis serogroup B
strain MC58
Science 287 (5459), 1809-1815 (2000)

JOURNAL
MEDLINE
PUBMED
AUTHORS
TITLE
2 (bases 1 to 10947)
Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
Dodson,R.J., Nelson,W.C., Gwin,M.L., DeBoy,R., Peterson,J.D.,
Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
Dougenly,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
Cittone,H., Clark,E.B., Cotton,M.D., Uterback,T.R., Khouri,H.,
Oin,H., Yamathayan,J., Gill,J., Scariato,V., Maignani,V.,
Piazza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
Rappoli,R. and Venter,J.C.
Direct Submission

JOURNAL
FEATURES
SOURCE
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
Location/Qualifiers
1..10947
/organism="Neisseria meningitidis MC58"
/strain="MC58"
/db_xref="taxon:122586"
/note="serogroup: B"
815..1634
/gene="NMB1636"
/note="This region contains an authentic frame shift and
is not the result of a sequencing artifact; similar to
PID:2367651 percent identity: 100.00; identified by
sequence similarity; putative; opacity protein, authentic
frameshift"
complement(1835..2392)
/gene="NMB1637"
complement(1835..2392)
/gene="NMB1637"
/note="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF41986.1"
/db_xref="GI:7226887"
/translation="MDALKLTTRRSKRLKHPAPDAELEIOFQATQVDDHGMRP
FRFVIGGVGVRQPRDYLKQYALNFGDPAKKAKEKVMAMVIGYTRAPNDVP
KRKPEQMLTRACCAIVADLAATPAQGFQDVMWITGMVNSPLREARFCADKDTIGL
MMVGTPEVHVRKNTDEAFVSHW"
complement(2468..4102)
/gene="NMB1638"
complement(2468..4102)
/gene="NMB1638"
/note="similar to GB:U14003 SP:P30845 GB:D14055 PID:216537
PID:536958 percent identity: 64.40; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="YnhX/YhjW/Yjdb family protein"
/protein_id="AAF41987.1"
/db_xref="GI:7226888"
/translation="MIRPNRPRKSSALIAPLISVSLVLTNAFAKVVLEHPNGT
GADIFLYMPVVLFLSNVFPVIALPFRKVLIPILLVISAAYSQEIFENFNS
MLNVLQTTAAESARLITPGYVLIIVCLGLPALAYIAVVKYKVKRELITRLVLA
VSFLCALGAMLODYDASEFRNKSQVTLIVSNFGAGVSKDKKRSNIPYTDLD
MAVQNRPAQSILRRFVLVYGETTRRANGINGTSQTPLLAARGDEIVNPGVRS
GTSRAHSPCKMSTDRDTPYDEIKAEHODNLDIVORAGYEVTWLENDGCGGCKV
PNTDVTSLNLPYCRNGECLDNLITKRFDEVLNKNDKDAVLIHTIGSHGPPYERYT
EAERFTPTCDTNEINKCYRAVLVNTYDNTVLYVDQFIDVIRKLENRDLESVVHY
SDHGSLENGYELHAAPIAAPSQTTHIPMVWFSKAFRQHGIDFQCLKKAENE
YSHDYFSTVLILMIOISNQTAKREMDIIAACRRPR"
complement(4176..4505)
/gene="NMB1639"
complement(4176..4505)
/gene="NMB1639"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF41988.1"
/db_xref="GI:7226889"
/translation="MSRCILKRLQTAASAVSAGVLEVVRLCCRRSLVFNPSVMLF
VPDCEVGAAGFQYEMCLPFQWQRLQTSNLAIDVYVMSAGRLNGNRPHSRFLA
LLGSVAL"
complement(4505..5611)
/gene="NMB1640"
complement(4505..5611)
/gene="NMB1640"
/note="similar to PID:2253231 PID:2253236 PID:2253251
percent identity: 100.00; identified by sequence
similarity; putative"

/note="Correia element; hmms hit to HMM Correia (1 - 62),
score: 78.71"
/label=Correia
562. .604
/note="Correia element; hmms hit to HMM Correia (114 -
156), score: 51.92"
/label=Correia
complement(667) .676)
/note="Core DNA uptake sequence: gccgcctgaa"
/label=DUS
complement(668. .1813)
/gene="dape"
complement(668. .1813)
/gene="dape"
/EC_number="3.5.1.18"
/note="NMA1730, dape, succinyl-diaminopimelate
desuccinylase, len: 381aa; similar to many eg. SW:P24176
(DAPE_ECOLI) dape, succinyl-diaminopimelate desuccinylase
from Escherichia coli (375 aa) fasta scores; E(): 0, 56.98
identity in 371 aa overlap. Contains Pfam match to entry
PF01546 Peptidase_M20."
/codon_start=1
/transl_table=11
/product="succinyl-diaminopimelate desuccinylase"
/protein_id="CAB84958.1"
/db_xref="GI:7380372"
/db_xref="SPTREMBL:Q9JTL0"
/translation="MTETQSLEKALISRPVTPDDRCOKLVERLYKIPAAEEL
HFGDKNIMLRGTGVKPVVCGAGHTVPVPAVEKMPSPPEPERGCRATYGRCAAD
KTSICFVACERFAVEHPDHOSIALILTSDEGDLDTGTVVDILKAGELIDYC
IVGEPTAYDKLMDIKNGRGLSGNLTAKGQGHVAPHILNPVHTFAALLLEUQ
EWMDEGNKTFEPTFQISNNGTGATNVPGLNVKFNFRSTESTEGALIKORVHA1
LKHGVQYDLQWSCGPFLTQAGKLTVDVARAIAEFCIGIEALSTGTGUSDGREFIKA
IAKELIEGDSNATTHQINENVRDLIDPKLSAVYEGILARLACNAV"
complement(836. .1786)
/gene="dape"
/note="Pfam match to entry PF01546 Peptidase_M20, score
252.50, E-value 3.7e-72"
complement(1863. .1872)
/note="Core DNA uptake sequence: gccgcctgaa"
/label=DUS
complement(1927. .2547)
/gene="NMA1731"
complement(1927. .2547)
/note="NMA1731, conserved hypothetical protein, len:
206aa; similar to SW:P44013 (Y552_HA8IN) hypothetical
protein from Haemophilus influenzae (207 aa) fasta scores;
E(): 0, 59.68 identity in 203 aa overlap."
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAB84959.1"
/db_xref="GI:7380373"
/db_xref="SPTREMBL:Q9JQZ3"
/translation="MLTPKSCDIFNIPPFQFSOLKRYQPSISIQIKADYENWOIWOQ
LIQVVAELGAPAPPIHERWCGNVOVRAHFFAYFYEQKNSAAILSLINRRLSY
SLDMHCYKADVSPIALPDYRNMLDNDTEKYASFDMHGAESERDYRTAAQONESDR
KLQNDDEFICIKHIERDDLGRQDVAKWIAETVEDLLPIYEACHGK"
complement(2551. .2558)
complement(2570. .2579)
/note="Core DNA uptake sequence: gccgcctgaa"
/label=DUS
complement(2581. .3090)
/gene="NMA1732"
complement(2581. .3090)
/gene="NMA1732"
/note="NMA1732, conserved hypothetical protein, len:
169aa; similar to SW:P43971 (Y243_HA8IN) hypothetical
protein from Haemophilus influenzae (172 aa) fasta scores;
E(): 0, 49.78 identity in 169 aa overlap."
/codon_start=1
/transl_table=11

/product="conserved hypothetical protein"
/protein_id="CAB84960.1"
/db_xref="GI:7380374"
/db_xref="SPTREMBL:Q9JTK9"
/translation="MNPETQSVTFAPAEPIEMLYACHGKVRRCGOVAMLSDYIAENG
NOLVLTIRQITRYFNVAAPLHHEDEENFPPLLOVAPQAOESVDELROHVSILGN
MDAVAAEFPAKLEADNAVYIPDAEAFKRFVAGDVHIAIEPLFDMCKMFIPEKLETFIG
EIMAAAAARRR"
complement(3185. .3736)
/gene="NMA1733"
complement(3185. .3736)
/gene="NMA1733"
/note="NMA1733, H.8 outer membrane protein, len: 183aa;
strongly similar to SW:P07212 (H8_NEIME) H.8 outer
membrane protein from Neisseria meningitidis (183 aa)
fasta scores; E(): 0, 96.78 identity in 183 aa overlap.
Contains Pfam match to entry PF00127 copper-bind, Copper
binding proteins, plastocyanin/azurin family and prosite
match to PS00196 Type-1 copper (blue) proteins signature."
/codon_start=1
/transl_table=11
/product="H.8 outer membrane protein"
/protein_id="CAB84961.1"
/db_xref="GI:7380375"
/db_xref="SWISS-PROT:P57025"
/translation="MKAVLALISAIVIGLACSOEPAPAAEATPAEAPASEAPAE
AAPADAAPAAAGCAATVESNDMMQFPTKDIQVSKAKEFTITLKHTGTPKTSMG
NIVIGKTEMDGIFKDCVGAADPVVPRDDARVVAHTKLIGGESSLTIDPAKLADG
EYRFACFTFPHGALMNKRVLTVD"
complement(3191. .3568)
/gene="NMA1733"
/note="Pfam match to entry PF00127 copper-bind, Copper
binding proteins, plastocyanin/azurin family, score
218.60, E-value 9.1e-62"
complement(33212. .33259)
/gene="NMA1733"
/note="PS00196 Type-1 copper (blue) proteins signature."
complement(4017. .4439)
/gene="NMA1734"
complement(4017. .4439)
/gene="NMA1734"
/note="NMA1734, unknown, len: 140aa"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB84962.1"
/db_xref="GI:7380376"
/db_xref="SPTREMBL:Q9JTK8"
/translation="MNLQDGRDALLSGLIKQAGQWRRLDAVKKLLPANLPHHPQT
ACIEDRLVLIANNMASRLKMTAPSVLPQLAGIDASIRSVSLVLPKPKPTNT
LHLSKALLESFGSAAVKLEKRHPLEALNMLVRRHGA"
4306. .4315
/note="Core DNA uptake sequence: gccgcctgaa"
/label=DUS
complement(4486. .4495)
/note="Core DNA uptake sequence: gccgcctgaa"
/label=DUS
4570. .7320
/gene="seca"
4570. .7320
/gene="seca"
/note="NMA1735, seca, preprotein translocase Seca subunit,
len: 916aa; similar to many eg. SW:P10408 (SECA_ECOLI)
preprotein translocase Seca subunit from Escherichia coli
(901 aa) fasta scores; E(): 0, 58.08 identity in 920 aa
overlap. Contains Pfam match to entry PF01043
SecA_protein, SecA protein, amino terminal region and
prosite match to PS01312 Protein seca signatures."
/codon_start=1
/transl_table=11
/product="preprotein translocase Seca subunit"
/protein_id="CAB84963.1"
/db_xref="GI:7380377"

SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 70423)
TITLE Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL Mus musculus, clone RP24-181A10
REFERENCE Unpublished
2 (bases 1 to 70423)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barua,N., Bastien,V., Boguslavsky,L., Bouckgatter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,D.S., Dodge,S., Fato,S.,
Ferrelle,P., FitzHugh,W., Gage,D., Galagan,J., Gargyda,S.,
Glade,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Illey,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Laroque,K.,
Lamazares,R., Landers,T., Lenoczky,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Margulis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPeeters,R., Meldrum,J.,
Menus,L., Mihova,T., Mienna,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Riebeck,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tophan,K., Travers,M., Travis,N., Trigillo,J., Vassiliou,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All reads were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L18257
Center clone name: 181_A_10

* NOTE: This record contains 90 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1
* 674: contig of 674 bp in length
* 675 774: gap of 100 bp
* 775 1472: contig of 698 bp in length
* 1473 1572: gap of 100 bp
* 1573 2265: contig of 693 bp in length
* 2266 2365: gap of 100 bp
* 2366 3060: contig of 695 bp in length
* 3061 3160: gap of 100 bp
* 3161 3834: contig of 674 bp in length
* 3835 3934: gap of 100 bp
* 3935 4612: contig of 678 bp in length
* 4613 4712: gap of 100 bp
* 4713 5394: contig of 682 bp in length
* 5395 5494: gap of 100 bp
* 5495 6171: contig of 677 bp in length
* 6172 6271: gap of 100 bp

* 6272 6954: contig of 683 bp in length
* 6955 7054: gap of 100 bp
* 7055 7749: contig of 695 bp in length
* 7750 7849: gap of 100 bp
* 7850 8538: contig of 689 bp in length
* 8539 8638: gap of 100 bp
* 8639 9325: contig of 687 bp in length
* 9326 9425: gap of 100 bp
* 9426 10132: contig of 707 bp in length
* 10133 10232: gap of 100 bp
* 10233 10931: contig of 699 bp in length
* 10932 11031: gap of 100 bp
* 11032 11689: contig of 658 bp in length
* 11690 11789: gap of 100 bp
* 11790 12463: contig of 674 bp in length
* 12464 12563: gap of 100 bp
* 12564 13228: contig of 665 bp in length
* 13229 13328: gap of 100 bp
* 13329 14013: contig of 685 bp in length
* 14014 14113: gap of 100 bp
* 14114 14800: contig of 687 bp in length
* 14801 14900: gap of 100 bp
* 14901 15600: contig of 700 bp in length
* 15601 15700: gap of 100 bp
* 15701 16397: contig of 697 bp in length
* 16398 16497: gap of 100 bp
* 16498 17157: contig of 660 bp in length
* 17158 17257: gap of 100 bp
* 17258 17944: contig of 687 bp in length
* 17945 18044: gap of 100 bp
* 18045 18717: contig of 673 bp in length
* 18718 18817: gap of 100 bp
* 18818 19518: contig of 701 bp in length
* 19519 19618: gap of 100 bp
* 19619 20285: contig of 667 bp in length
* 20286 20385: gap of 100 bp
* 20386 21060: contig of 675 bp in length
* 21061 21160: gap of 100 bp
* 21161 21835: contig of 675 bp in length
* 21836 21935: gap of 100 bp
* 21936 22604: contig of 669 bp in length
* 22605 22704: gap of 100 bp
* 22705 23367: contig of 663 bp in length
* 23368 23467: gap of 100 bp
* 23468 24145: contig of 678 bp in length
* 24146 24245: gap of 100 bp
* 24246 24939: contig of 694 bp in length
* 24940 25039: gap of 100 bp
* 25040 25722: contig of 683 bp in length
* 25723 25822: gap of 100 bp
* 25823 26488: contig of 666 bp in length
* 26489 26588: gap of 100 bp
* 26589 27282: contig of 694 bp in length
* 27283 27382: gap of 100 bp
* 27383 28075: contig of 693 bp in length
* 28076 28175: gap of 100 bp
* 28176 28851: contig of 676 bp in length
* 28852 28951: gap of 100 bp
* 28952 29608: contig of 657 bp in length
* 29609 29708: gap of 100 bp
* 29709 30386: contig of 678 bp in length
* 30387 30486: gap of 100 bp
* 30487 31159: contig of 673 bp in length
* 31160 31259: gap of 100 bp
* 31260 31949: contig of 690 bp in length
* 31950 32049: gap of 100 bp
* 32050 32752: contig of 703 bp in length
* 32753 32852: gap of 100 bp
* 32853 33549: contig of 697 bp in length
* 33550 33649: gap of 100 bp
* 33650 34260: contig of 611 bp in length
* 34261 34360: gap of 100 bp
* 34361 34999: contig of 639 bp in length

```

* 35000 35099: gap of 100 bp
* 35100 35801: contig of 702 bp in length
* 35802 35901: gap of 100 bp
* 35902 35993: contig of 694 bp in length
* 36596 36695: gap of 100 bp
* 36696 37401: contig of 706 bp in length
* 37402 37501: gap of 100 bp
* 37502 38182: contig of 681 bp in length
* 38183 38282: gap of 100 bp
* 38283 38951: contig of 669 bp in length
* 38952 39051: gap of 100 bp
* 39052 39717: contig of 666 bp in length
* 39718 39817: gap of 100 bp
* 39818 40501: contig of 684 bp in length
* 40502 40601: gap of 100 bp
* 40602 41287: contig of 686 bp in length
* 41288 41387: gap of 100 bp
* 41388 42086: contig of 699 bp in length
* 42087 42186: gap of 100 bp
* 42187 42884: contig of 698 bp in length
* 42885 42984: gap of 100 bp
* 42985 43632: contig of 648 bp in length
* 43633 43732: gap of 100 bp
* 43733 44422: contig of 690 bp in length
* 44423 44522: gap of 100 bp
* 44523 45221: contig of 699 bp in length
* 45222 45321: gap of 100 bp
* 45322 46018: contig of 697 bp in length
* 46019 46118: gap of 100 bp
* 46119 46780: contig of 662 bp in length
* 46781 46880: gap of 100 bp
* 46881 47539: contig of 659 bp in length
* 47540 47639: gap of 100 bp
* 47640 48318: contig of 679 bp in length
* 48319 48418: gap of 100 bp
* 48419 49095: contig of 677 bp in length
* 49096 49195: gap of 100 bp
* 49196 49883: contig of 688 bp in length
* 49884 49983: gap of 100 bp
* 49984 50679: contig of 696 bp in length
* 50680 50779: gap of 100 bp
* 50780 51488: contig of 709 bp in length
* 51489 51588: gap of 100 bp
* 51589 52278: contig of 690 bp in length
* 52279 52378: gap of 100 bp
* 52379 53051: contig of 673 bp in length
* 53052 53151: gap of 100 bp
* 53152 53848: contig of 697 bp in length
* 53849 53948: gap of 100 bp

```

Query Match 2.5%: Score 42.2; DB 2; Length 70423;
 Best Local Similarity 36.3%; Pred. No. 15; Matches 136; Conservative 0; Mismatches 238; Indels 1; Gaps 1;

```

OY 1217 caaacaacgcccgaacccaagtcgcttgaaggltaagggllccgaatttg 1276
DB 65100 CAGAAGTACGCAAACTTGTGACTAGCTGTTCTAGGCTGCTATGCTATGTTATGTTAAT 65041

OY 1277 aaaaagaagtaaatcagatcagaagaataatccgcctgacacaaagttaattctatag 1336
DB 65040 ACATGCAAGCAAAACAAAGCAAGAGAGATGAATTCACACGCCAGTGGCGATGATTC 64981

OY 1337 atgaaccgcttlaattcctaagaagttctgctgcatcattcttgcgtactaactg 1396
DB 64980 AGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 64921

OY 1397 ccagaattcaatcagcaaatltaaccaagcagtagaatcagatatatccacataaaa 1456
DB 64920 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 64861

OY 1457 attactctccctcagcagcgtcacaagaagacct-aataatgatatattggataaattc 1515
DB 64860 GTTATATGCTGTATGATCAATAATAAAGCTTATGAGATGATGAGCGGAATATGTGA 64801

```

```

OY 1516 gtaaatgactgaatgctccatcaagaactaaagtcagaatttgatgagatgt 1575
DB 64800 GGTATTATGAGGAGAGAAATGAAGGTATGAAATTTTCCCAATTCAGTGTACTGATGTATG 64741

OY 1576 caattgtctaaaca 1590
DB 64740 AAATATCAAAAATA 64726

```

RESULT 45
 LOCUS AP004215
 DEFINITION Oryza sativa chromosome 7 clone OJ1429_G07, *** SEQUENCING IN
 PROGRESS ***, in ordered pieces.

ACCESSION AP004215
 VERSION AP004215.1 GI:15787872
 KEYWORDS HTGS-PHASE2.
 SOURCE Oryza sativa (cultivar:Nippondare) DNA, clone:OJ1429_G07.
 ORGANISM Oryza sativa

REFERENCE
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
 TITLE Oryza sativa nippondare(GA3) genomic DNA, chromosome 7, BAC
 clone:OJ1429_G07
 JOURNAL Published Only in Database (2001) In press
 REFERENCE 2 (bases 1 to 98980)
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
 TITLE Direct Submission
 JOURNAL Submitted (26-SEP-2001) Takuji Sasaki, National Institute of
 Agrobiological Sciences, Rice Genome Research Program; Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
 Tel:81-298-38-7441, Fax:81-298-38-7468)
 The nucleotide sequence of this BAC clone was generated by
 combining Monsanto and RGP-Japan sequencing data.

NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
 source
 1..98980
 /organism="Oryza sativa"
 /cultivar="Nippondare"
 /db_xref="taxon:4530"
 /chromosome="7"
 /clone="OJ1429_G07"

BASE COUNT 28747 a 20334 c 20154 g 29695 t 50 others
 ORIGIN

Query Match 2.5%: Score 42.2; DB 2; Length 98980;
 Best Local Similarity 49.3%; Pred. No. 16; Matches 110; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

```

OY 1449 acctaaatctctctcctcctcagcagcgtcacaagaagcccaataatgatatgga 1508
DB 86070 ACTTACATGAGATTTCATCATGAGCGCATRACCAACCAAAAACATRAAGTATCTGCT 86129

OY 1509 taattgttaatgaatgactaaagtcacataagaactaaagtcagaattggaatg 1568
DB 86130 TAAAGTTGAATTAACATGATTAAACACCTTAAGTCTAAATATACATATAGCTTAAT 86189

OY 1569 ggaatgtcaatgtctcaaaacaggaagagacaaacttgatggcctagtggaatgtaa 1628

```

Db 86190 GTTCGATTGGCAATTAGAAATGTCAGTCTAAGACTTGTGCTAGATTAGATGTTA 86249
Oy 1629 gcatttaaatatataatgatggaagaattacacacaalga 1671
|| ||| | || ||||| || || || ||
Db 86250 GCCAGATTATACACTTATATTGGAAAAATGAACTCTAAAGA 86292

Search completed: June 30, 2002, 09:25:29
Job time: 58147 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2002, 07:52:06 ; Search time 27321.5 Seconds
(without alignments)
1445.323 Million cell updates/sec

Title: US-09-303-518D-463

Perfect score: 1887
Sequence: 1 ttgggcattccccaataa.....catatccattatgaatag 1887

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 100 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_bg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1886	99.9	1887	6	A96430	A96430 Sequence 46
2	1427.8	75.7	1827	6	AX044009	AX044009 Sequence
3	1427.8	75.7	11796	1	AE002419	AE002419 Neisseria
4	1427.8	75.7	349980	6	AX044030	AX044030 Sequence
5	1231.2	65.2	1671	6	A96432	A96432 Sequence 46
6	1231.2	65.2	311321	1	NMA372491	AL162754 Neisseria
7	1198.8	63.5	2304	6	AX236465	AX236465 Sequence
8	1198.8	63.4	2028	6	AX236421	AX236421 Sequence
9	1196	63.4	2028	6	AX239731	AX239731 Sequence
10	1196	63.4	4425	6	AX236407	AX236407 Sequence
11	1196	63.4	4425	6	AX239715	AX239715 Sequence
12	1194	63.3	2019	6	AX236447	AX236447 Sequence
13	1194	63.3	2019	6	AX239733	AX239733 Sequence
14	1194	63.3	2256	6	AX236451	AX236451 Sequence
15	1194	63.3	2256	6	AX236459	AX236459 Sequence
16	1194	63.3	2256	6	AX239737	AX239737 Sequence
17	1194	63.3	2256	6	AX239745	AX239745 Sequence
18	1194	63.3	2421	6	AX236449	AX236449 Sequence
19	1194	63.3	2421	6	AX236453	AX236453 Sequence
20	1194	63.3	2421	6	AX239735	AX239735 Sequence
21	1194	63.3	2421	6	AX239739	AX239739 Sequence
22	1194	34.1	683	6	A96426	A96426 Sequence 45
23	642.8	17.1	381	6	A96424	A96424 Sequence 45
24	322.4	15.0	11700	1	AE002393	AE002393 Neisseria
25	283.4	15.0	349980	6	AX044029	AX044029 Sequence
26	283.4	14.3	195767	1	NMA722491	AL162752 Neisseria
27	269.6	12.0	10057	1	AE002420	AE002420 Neisseria
28	226.6	6.3	1497	6	A96138	A96138 Sequence 17
29	119.8	6.3	1744	6	A61829	A61829 Sequence 9
30	119.8	6.3	3287	6	A61821	A61821 Sequence 1
31	119.8	6.3	3294	1	AF142582	AF142582 Neisseria
32	119.8	6.3	3294	1	AF142582	AF142582 Neisseria
33	108.2	5.7	340806	1	NMA122491	AL162752 Neisseria
34	105.2	5.6	1449	6	A96134	A96134 Sequence 16
35	105	5.6	1452	6	A96132	A96132 Sequence 16
36	105	5.6	10869	1	AE002559	AE002559 Neisseria
37	105	5.6	172325	6	AX044035	AX044035 Sequence
38	96	5.1	375	6	A96130	AX044035 Sequence 16
39	80.2	4.3	7218	6	166494	A96130 Sequence 16
40	49	2.6	125020	9	AF429315	A96130 Sequence 14
41	48.4	2.6	125020	9	AF429315	AF429315 Homo sapi
42	45.8	2.4	41755	2	AC006179	AF429315 Homo sapi
43	45.2	2.4	115890	2	AR135147	AC006179 Homo sapi
44	45.2	2.4	115890	2	AP002824	AR135147 Sequence
45	45.2	2.4	157078	1	SC4C6	AP002824 Homo sapi
46	44.6	2.4	30941	1	AE006028	AC096727 Homo sapi
47	44.2	2.3	10598	1	AE006028	AL079355 Streptomy
48	44.2	2.3	12026	1	AE005109	AE006028 Caulobact
49	44.2	2.3	169022	2	AC027354	AE005109 Halobacte
50	43.8	2.3	605	8	AF377524	AC027354 Homo sapi
51	43.8	2.3	605	8	AF377525	AF377524 Zea mays
52	43.8	2.3	605	8	AF377527	AF377525 Zea mays
53	43.8	2.3	605	8	AF377536	AF377527 Zea mays
54	43.8	2.3	605	8	AF377542	AF377536 Zea mays
55	43.6	2.3	291	3	AF134687	AF377542 Zea mays
56	43.6	2.3	204652	2	PFMAL13P6	AF134687 Plasmodiu
57	43.4	2.3	231	3	AF134664	AL049183 Plasmodiu
58	43.4	2.3	131535	9	AC005875	AF134664 Plasmodiu
59	43.4	2.3	133954	9	AL355493	AC005875 c1tb_188-
60	43.4	2.3	204881	2	AL591111	AL355493 Human DNA
61	43.2	2.3	674	3	PEFA310620	AL591111 Homo sapi
62	43.2	2.3	674	3	PEFA310621	AJ310620 Plasmodiu
63	43.2	2.3	674	3	PEFA310623	AJ310621 Plasmodiu
64	43.2	2.3	347660	1	AP002994	AJ310623 Plasmodiu
65	43	2.3	942	1	AF169562	AP002994 Mesorhizo
66	42.6	2.3	3151	1	AB055506	AF169562 Mesorhizo
67	42.4	2.2	15311	1	AE007164	AB055506 Streptomy
68	42.2	2.2	605	8	AF377520	AE007164 Mycobacte
69	42.2	2.2	605	8	AF377521	AF377520 Zea mays
70	42.2	2.2	605	8	AF377522	AF377521 Zea mays
71	42.2	2.2	605	8	AF377523	AF377522 Zea mays
72	42.2	2.2	605	8	AF377526	AF377523 Zea mays
73	42.2	2.2	605	8	AF377529	AF377526 Zea mays

Db	1261	GGGTTCCGAATTTTGAGAACGACGTGAATATGATACGAACTCGATATTTCAGAAATTA	1320
Oy	1321	tcggtggcggtatcctaacgaagcctgctgtttgtagcgaaccgagatgggaagt	1380
Db	1321	TCGGGGGGCGGTATACCTAAGGGCTAAGCCTGTGTTGATGGGAAACCGAGATGGAGGTT	1380
Oy	1381	gataagaagcttaataatgacacactcgtgagcaggtggggaataatgcttaggaagac	1440
Db	1381	GATAGGAGCTTAATAAATGACAACTGTAAGGTGGAGGAAATAATTTAGCAAGACG	1440
Oy	1441	agaagaagagctcagagtagctcaagtttaaaagcccatcgcaacgagaatgggaaataaa	1500
Db	1441	AGAAGAAGAGTACAGTAGTACAGTTTAAAGCCCATGGCGAACGAAATGGGAAATATA	1500
Oy	1501	acaaggttagatlttaacatcttaataagtggtgatcatcaataagaagaagcagtaaca	1560
Db	1501	ACAGGGTTAGATTTAATTCATTTATAGTGTGTCATATCAATTAAGAAAGACAGATACA	1560
Oy	1561	ggagggcatagcttaaccggtggtgtagttagcgggtgatacaaaacctcgacactgat	1620
Db	1561	GGAGGGCATAGCTTAACCCGTGTGATGTAGCGGTATACAAACAACTTCGGCACCTGAT	1620
Oy	1621	aaacatgggtctatcaacgcacagctggaaatlaaaagccctgagtgaagtgggaagt	1680
Db	1621	AAACATGGGCTTTATACAGCCAGCAGTGAATTAATAAACCTGATGAGATTGGAGGTG	1680
Oy	1681	aaaaaagaaagtggaagatgtagcaacacacacatgttcccaaaagattggat	1740
Db	1681	AAAAAGAAAGAGTGGGGAAGTGAACCAAGCACACCATGTTCCCAAAAGATTGGAT	1740
Oy	1741	gaagctagaatagggcgtgaagttacttcggtctggaaagtgaataatgcttagaagt	1800
Db	1741	GAGGCTAATAATTAGGGCTGAGACTTACTTCGGCTGGGGAAGTAGAATGCTTAAGAT	1800
Oy	1801	aataatgacaggtgacaaagtaacatcggtgtatlaaataagaagattaccgaacctaat	1860
Db	1801	AATAATGACAGGCTACAAAGTAATCGGCTATTAAATAGAAAGATTACCGAACCTAAT	1860
Oy	1861	agaacagcatatcccatctttagaataag	1887
Db	1861	AGAACAGCATATCCCATTTATGATAAG	1887
RESULT 2			
AX044009		1827 bp	DNA
LOCUS	AX044009	1827 bp	DNA
DEFINITION	Sequence 88 from Patent WO0066791.		linear
ACCESSION	AX044009		PAT 24-NOV-2000
VERSION	AX044009.1		GI:11342894
KEYWORDS			
SOURCE			
ORGANISM	Neisseria meningitidis.		
	Neisseria meningitidis		
	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;		
	Neisseria.		
REFERENCE	1 (bases 1 to 1827)		
AUTHORS	Pizza, M., Hickey, E., Peterson, J., Tettein, H., Venter, J. C.,		
	Masigam, V., Galeotti, C., Mora, M., Ratti, G., Sarselli, M.,		
	Scalafco, V., Rappelli, R., Frazer, C. M. and Grandi, G.		
TITLE	Neisseria genomic sequences and methods of their use		
JOURNAL	Patent: WO 0066791-A 88 09-NOV-2000;		
FEATURES	CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)		
source	1. 1827		
	/organism="Neisseria meningitidis"		
	/db_xref="taxon:487"		
BASE COUNT	539 a 446 c 454 g 387 t	1 others	
ORIGIN			
Query Match	75.7%;	Score 1427.8;	DB 6; Length 1827;
Best Local Similarity	95.68;	Pred. NO. 0;	
Matches 1468; Conservative	0;	Mismatches 68;	Indels 0; Gaps 0;

Oy	1	ttggacattcccgcaaaataatcccttatctctgcalactgycagtgctcgcatg	60
Db	1	TTGGGCAATTTCCGGAAATATATCCCTTATCTGTGCATACGTGCAGTGTGCTGCCGATG	60
Oy	61	catgacaagcctcagaattggcaaaagatcccttatccggcaggtctcagcctga	120
Db	61	CATGCAACGCCCTCAGATTGGCAAAAGATTCTTTATCCGGCAGGTTCTGACCGTCA	120
Oy	121	cattgaaacccgacgggaaataccacctatctcgcaacagggaggtctgacnagcg	180
Db	121	CATTTCACACCCGACGGGAAATACCACCTATTCCGACACAGGGGGAATCTGCCAGGCG	180
Oy	181	aacggccataatcgatttgggaaacatacaaaagccatcagttgggcacctgatatca	240
Db	181	AGCGGCATATCGGATTGGGAAATAACAAAGCCATCGTTGGCCACACTGATGATTCAA	240
Oy	241	caagggcctgtgagaagaataatcggtcaacttgcgcctttccgatacagggacaa	300
Db	241	CAGCGGCCCATTTAAAGGAATATCGGCTACATTTGCGCTTTCCGATACGGGACGAA	300
Oy	301	ttccattcgcccttcgaaacacatgacctcacatctcgatctcgaagccgtagtccc	360
Db	301	GTCATTCCCTTCCTTCGACAAACCATGCTCACATTCCGATTCGATGAGAGCCGGTAGTCC	360
Oy	361	gttagcagatcaagccttaccgcatccattgggaacgagatacgaacacatcccgagac	420
Db	361	GTTACAGGATTATAGCTTTACCGCATTCATTGGAGGATACGAACACCATCCCGCCGAC	420
Oy	421	ggtcatgacgggccaacagggcggtctatcccgctcccaaaagggcgagagatatag	480
Db	421	GGCTATGACGGGCCCAAGGCGGGCGGCTATCCGCTCCCAAGGCGCGGATATATAC	480
Oy	481	agctacgacataaaagggcgttgcccaaaataatccgctcaactcagacgacacgacg	540
Db	481	AGCTACGACATAAAAGGGGTTGCCAAATATATCCGCTCAACCTGACCGACACCGCAGC	540
Oy	541	accggaacacggcgttcgcgaaccgttccaagaatgcgcgcgtatgctgacgcaagagta	600
Db	541	ACCGGACAAAGGCTTGCGGACCGCTTCCCAATGCGGATGATGATGTCGACGCAAGAGTA	600
Oy	601	ggcgaagatccaacgagccgacccgatacagcccgagctgagacatcgagcaatgccc	660
Db	601	GGCGACGATTCMAAGCGGCCACCGGATACGCCCGGACTGGACAGATCGGCAATGCC	660
Oy	661	gccgaagccttcaacggcactgacagataatcgtcaaaacatcgcgcgcgagagagaa	720
Db	661	GCCGAAGCCTTCAAGGCACTGACAGATATCGTTAAACATATCGCGCGGCGCAGAGAA	720
Oy	721	attgctgcgcagcgatgctgcaagggtataagcgaagctcaaacattgctgcatg	780
Db	721	ATTGTCGGGCGAGCGGATGCCGTGCGAGGCAATAAAGGCTCAAAACATTTGCTCATG	780
Oy	781	caagccttggtgctgcttccacccgaaacagaatggcgacatcaacgaattggcgat	840
Db	781	CACGGCTTGGGTCCTCTTCCACCGGAAACAAATGCGCGCATCAACGATTGGCGCAT	840
Oy	841	atggcgcaactcaaaagactatccgcagcagccatccgcatctggcgaagtcacaaaccc	900
Db	841	ATGGCGCAACTCAAAAGACTATCGCGCAGCAGCATCCGCGATTGGGAGTCCAAACCC	900
Oy	901	aatgcgcgaacagacatagaagcgtcagcaaatatctttatggcagcattccccatcaa	960
Db	901	AATGCCGCAACAGGCAATGAAGCGCCGCAAAATCTTTATGSCAGGCATCCCATATA	960
Oy	961	gggattgagcctgctccgggaaataacggtctggcgacatcaacgacatctctgacaag	1020
Db	961	GGGATTGAGCTGTTTCGGGAAATATACGGTGTGGGGGATCAAGGCACATCTATCAAG	1020
Oy	1021	cggtcgcaatggcgcgatcgcatctgcgaagaaggaatccgcgtcgaacgacatcttc	1080
Db	1021	CGGTGCAATGGCGCGATCGCATCTGCCGAAAGGGGAAATCCCGGTCACGCAATTTT	1080
Oy	1081	gccgatgcgcatatgcgaataatcccgctcccttacaattcccgaaatattcgttcaac	1140

Db	1081		source
Qy	1141	ttggagcagcgtttacggcaagaacaatcacctcctcaaccgttgcgcgtcaacgac	1. .11796
Db	1141	TTGGAGAGAGGTTACGCGCAAAAGAAATCATCCTCTTCACACCGTGGCGGTCAACAGCGC	/organism="Neisseria meningitidis MC58"
Qy	1201	aaaatgtcaaatcttgcagacccacgcccaggaagcaggttaccgtttgcagcgttaaa	/strain="MC58"
Db	1201	AAAAATGTCAAACTGTGCGACACCAACGCCACCGAAGACAGCGCTACCGTTTGACGCGTAAA	/db_xref="taxon:122586"
Qy	1261	gggttccgaatttcgaaagcagcgtgaatatgatcagaagctcgatattcaagaatta	/note="serogroup: B"
Db	1261	GGGTTTCCGAATTTTGAAGACGACGTGAATATGATACGAAGCTCGATATTCAGAAATTA	complement(174..1139)
Qy	1321	tcggggggcgggttacttaagcgttaagccgtttttgatgcgaaccgagatggaggtt	/gene="NMB0635"
Db	1321	TCGGGGGGCGGTTACTTAAGCCTTAAGCCGTGTGTTGATGCGAACCAGATGCGAGGTT	complement(174..1139)
Qy	1381	gataagaaacctataaatgtacaactcgtgagcaggttgagaanaattgtcaggaagc	/gene="NMB0635"
Db	1381	GATAGGAACCTTAAATTAATGTACACTCGTGACAGCTGGAGAAAATGTCTCAGGAAATA	/note="similar to SP:p37247 PID:1197006 percent identity: 61.94; identified by sequence similarity; putative"
Qy	1441	agaagaagcagtgatgagtgatgcgttttaagaagcccatgcgcgaacgagaatgggaaataa	/codon_start=1
Db	1441	AGGAACGGTAATTAATAACGTAATCTTACCAACATGCTCACTAGAGAGGAAATTAAT	/transl_table=11
Qy	1501	acagaggttagatttaattcaattatgaagttgagttgag	/product="transposase, IS30 family"
Db	1501	AAACTAAATCTGCGCATGATGAATTAATTTTGCAGAT	/protein_id="AAFA1058.1"
RESULT	3		/db_xref="GI:7225864"
LOCUS	AE002419	11796 bp	/translation="MSYTLQNGERYHTIOLSRHCTVETIAIKOLNRHSTISRTIRRR
DEFINITION	Neisseria meningitidis serogroup B strain MC58 section 61 of 206 of the complete genome.		RTGGQVYAEKARQSRFTIKRQRPYLDLSOLIOHIDPLTRKISPEQVAYLCKHH
ACCESSION	AE002419	AE002098	RTGGQVYAEKARQSRFTIKRQRPYLDLSOLIOHIDPLTRKISPEQVAYLCKHH
VERSION	AE002419.1	GI:7225863	QTLHSTRIVLRQDKNGSTLMQHLATICKSPTRKRGSWTGKPNVNRGIERPRP
KEYWORDS			IYDOKSRIGDNEADITVGKGSALLTVEVTRTTIICKLDSLKAEPTAAAVALK
SOURCE	Neisseria meningitidis MC58.		AHKDVHTITMDNGKEFYQHKIKALKAEFTYPCRPYRSMKGLNENTNGLIRQYFPFP
ORGANISM	Neisseria meningitidis MC58		QTFDENISDRIRIROYDELNHRPRKLTGEPVSLFNLFPQLIH"
REFERENCE	1 (bases 1 to 11796)		complement(1303..1686)
AUTHORS	Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E., Olin,H., Yamaharan,J., Gill,J., Scarlato,V., Massignani,V., Rappapoli,R. and Venter,J.C.		/gene="NMB0636"
TITLE	Complete genome sequence of Neisseria meningitidis serogroup B strain MC58		complement(1303..1686)
JOURNAL	Science	287 (5459), 1809-1815	/gene="NMB0636"
MEDLINE	20175755		/note="hypothetical protein; identified by glimmer2; putative"
PUBMED	10710307		/codon_start=1
REFERENCE	2 (bases 1 to 11796)		/transl_table=11
AUTHORS	Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C., Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F., Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D., Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D., Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E., Olin,H., Clark,E.B., Cotton,M.D., Uitterlinden,T.R., Khouri,H., Pizsa,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R., Rappapoli,R. and Venter,J.C.		/product="hypothetical protein"
TITLE	Submitted (17-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA		/protein_id="AAFA1059.1"
JOURNAL			/db_xref="GI:7225865"
FEATURES	Location/Qualifiers		/translation="MTDTIDQDFEQQMMQAVDKLFEGHQDKLESMDGRQELVLTW
			RTEADIGSGTLOFVCMWGPFAAEKGTQSVLKKIGAVHSAMLHRAADALDKETRLOS
			EGKNLKEWMDITSTRQNRLEPAAQDSG"
			complement(11744..3120)
			/gene="NMB0637"
			complement(11744..3120)
			/gene="NMB0637"
			/note="similar to PID:1652362 percent identity: 67.97; identified by sequence similarity; putative"
			/codon_start=1
			/transl_table=11
			/product="argininosuccinate lyase"
			/protein_id="AAFA1060.1"
			/db_xref="GI:7225866"
			/translation="MHKRTWSGRNPEVSELVKQYASTIGFDRRLAEMDIQSLAQAQ
			MLKETGVLGDSGLADIIRGMAELIEETRISCKIEMSSDLEVDYHNHIERRLDKIGDGK
			RLHTRGRNQDVATDIRLMARDQITTVTOSLQSLQALDLACQNAETVAPGFTNHLQ
			AQPVSEGHMLAVEMIGRNERMADRCRVNNMIPGAALAGTTPYIOETREIATLIG
			FEQICONSILAVSDRDAIEPTAAASLVVHLSRLSEELILMSSPRGFGDIDRQCT
			GSSIMPKRNPDPVETLVKSGSGVIGLITLILMSQPLATKYKNDQEKEPLEPTAD
			TLDLRLRYADMMKRGVTVKRDNRRAVMQGFATATDADTLVKGAFPRAHSHVAQA
			VRHDAQVDLSELPLEVLQGFSDLLADVDYGVLTPEGSILANRHLGTAPEQVPRQV
			KRMKEMLA"
			complement(3139..4008)
			/gene="NMB0638"
			complement(3139..4008)
			/gene="NMB0638"
			/note="similar to PID:1628575 percent identity: 74.13; identified by sequence similarity; putative"
			/codon_start=1
			/transl_table=11
			/product="urp-glucose-1-phosphate uridylyltransferase"
			/protein_id="AAFA1061.1"
			/db_xref="GI:7225867"
			/translation="MKPIKRAVFPVAGMGCTRPLPAKSPKEMLPIDVRLQYAVEE
			AVAGCTEAVFVGNRKRSEIDHFDAVELELELEKRRHDKLEHVRNLIPNITCLY
			IRQAEALGIGHAVLCARAAIGDEPFAVITLADLLIDAPKALCOMVTVGRSGNSITG
			ETVEASQTSYSGIVTEOLKORPRTGIYKREKDPAPSNLAVGRTYILTPPIFDLT
			NLRGAGNEIOLTDGIAKLIDHEFVLAAHFPEGTRDCKSGKLGLETVAVGIAKHPTG
			PPPRRLIKKVRP"

```

1. .11796
/organism="Neisseria meningitidis MC58"
/strain="MC58"
/db_xref="taxon:122586"
/note="serogroup: B"
/complement(1174. .1139)
/gene="NMB0635"
/complement(1174. .1139)
/gene="NMB0635"
/note="similar to SP:P37247 PID:1197006 percent identity: 61.94; identified by sequence similarity; putative"
/codon_start=1
/product="transposase, IS30 family"
/protein_id="AA041058.1"
/db_xref="GI:7225864"
/translation="MSTQQLTQGERHYIQLSRHCVTEYELAKQLNRHKSTISREIRRI
RTGGQDSIAKARQDSRTIKQRKQPKLDSQLIQHLDPLRKLSPQYCAVLCYKHQ
QIHLHSTYATRYROKDSNSGTQMHQRLICSPYRKRYGWTGKGVNRYGIRNRP
IYQDSSTIGMEADTYIKQSKSLALTYLVREYTYTICKLSDKAEDTAAVRAALK
AHDQRYHTIMQDNGEYQHTYTKLTKAEYFCORYPSHMKSGIENMTNGLIRQYFPR
QDPERNSDEIRYODELNHRKRTLGTEYPSVLFNLQPLIH"
/complement(1103. .1686)
/gene="NMB0636"
/complement(1103. .1686)
/note="hypothetical protein; identified by glimmer2; putative"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AA041059.1"
/db_xref="GI:7225865"
/translation="MNDTPQADRFQPMQAVDKLFEQHDGKLESMDGREGQELVILW
READNGNSITLQFVQNPQAPAEKTCVLTAKGVASHNLIRADALDKETRRLOS
EKKNLKEMDITRSROONRLTAQSG"
/complement(11744. .3120)
/gene="NMB0637"
/complement(11744. .3120)
/note="similar to PID:1652362 percent identity: 67.97; identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="argininosuccinate lyase"
/protein_id="AA041060.1"
/db_xref="GI:7225866"
/translation="MMDKRWGSRFNPSEVELKQYASTIGFDRILAEWIDQSLANQ
MKEGVLDDEGLDARLQWAELEIRSGKLEWSSDLVDYHNIRRLRDLKGDGK
RLHTRSNNQDVAOTRIQLMRLQIYVLSIOSLQYALDLDAQNAETWPPGTHIDQ
AQPVSFGHMLAYVEMLGDNREMDACRVRNMRPLGALAGTTPPIQETIAELG
FEQIQNSIDAASDDQFALFETPAASLVNHLRSSEPLIWNMSPRGFGLIDARCTQ
GSSIMQPKNPVPELVYKSPRDMRAVMOGFATVDLADYLVKGMGPFRAHNVVAQ
VLIDILRIYADAMPRTYKSPRDMRAVMOGFATVDLADYLVKGMGPFRAHNVVAQ
VHRAQACVDSLELFLVYLGFSGLDADVDYGVLPREGSLNRNHLGGTAPAEVRFQV
KWRKMLA"
/complement(3139. .4008)
/gene="NMB0638"
/complement(3139. .4008)
/note="similar to PID:1628575 percent identity: 74.13; identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="UTP--glucose-1-phosphate uridylyltransferase"
/protein_id="AA041061.1"
/db_xref="GI:7225867"
/translation="MKPRKAVFPAVAGGTPLPPTKASPKRMLPIVQKPLQYNEE
AVEGCEPMVTVYKNSKSIDNHPKALELELEMRKDKLLEHYRNLIPNTICY
IROMALGIGAVNLACARALIDEPFAYLADLDLPARGALKQWQEVGSRNSNLSIG
ETVVASQSGSIGVETEDLKQFQRTTGIVKPKREDASNLAVGRTYITPPIFDLIT
LHPGAGNGEITDGLIAKLIDHEFLAHPFEGRTYDCSGKIGLEATYVGLKHPETG
PPRRLLKEVTE"

```

gene complement(4036. .4635)
/gene="NMB0639"
/complement(4036. .4635)
/gene="NMB0639"
/note="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAFA1062.1"
/db_xref="GI:7225868"
/translation="MSEKPEKIVLASGNACKLEFGNMFPRYSITVLPOSAGFIPCEP
EYPTFVENALAKARAKYSGLPALADDSICAAALGACISARAGNPNPSDA
NNLKLAELVGKADKSCGYCVLVFVRHKDDPPRIIEGVMWGMNDTPIGQNEFGVD
PYFYLPHEKTKTAELDTEVKNRESHRAQALAEILRLAL"
/complement(4628. .4861)
/gene="NMB0640"
/complement(4628. .4861)
/gene="NMB0640"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAFA1063.1"
/db_xref="GI:7225869"
/translation="MNRKTYLSVALFTIAFMLVLGLVGLTVGSKAFVASFRA
FGALFQIGSLALYLNKSLDRAQASATKENRY"
/complement(4969. .5502)
/gene="NMB0641"
/complement(4969. .5502)
/note="similar to GB:J42023 SP:P44529 PID:J003152
PID:J222038 PID:J204383 percent identity: 92.61;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="inorganic pyrophosphatase"
/protein_id="AAFA1064.1"
/db_xref="GI:7225870"
/translation="MADFQIILTPGDVGIIINVNEIPAGSNIRKIEWRKLAFOLD
RVEPAIFAKPTNNGFIPTQLEDGEDLVLLVEOPLATGVLEARVIGWKFVDDGE
VDDKIVCVADDDNNNGNAKYTSLDLPQOLIKQIEFPHNHKDLKKAGTTKVESGDAA
EAKKVIKESIERNKOA"
/complement(5603. .6061)
/gene="NMB0642"
/complement(5603. .6061)
/note="similar to GB:D10165 SP:P24236 PID:S81152
PID:669113 PID:912430 percent identity: 58.45; identified
by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="dATP pyrophosphohydrolase"
/protein_id="AAFA1065.1"
/db_xref="GI:7225871"
/translation="MMAKPKLPVSAIVLYSGDGGIILIRTPHSGFMQSYNGSL
PGEIVAQTRAREVMEETGILLADGQDDMDSTYELTHMRHRYPGGVENREHLFS
AEIPRDTALQDEEHVSTGWFMEAEAEVFSNRRALILEGRFGKR"
/complement(6348. .6772)
/gene="tRNA-Pro-3"
/product="tRNA-Pro"
/complement(6348. .6772)
/gene="tRNA-Pro-3"
/complement(6348. .6772)
/gene="NMB0643"
/complement(6348. .6772)
/note="MaFb-related protein; identified by Oxford;
putative"
/codon_start=1
/transl_table=11
/product="MaFb-related protein"

/protein_id="AAF62313.1"
/db_xref="GI:7413435"
/translation="MGISRRISLIISLIIVLCUPMHAASDLSDFIROYLIDROHPEP
DGYHILFGSRGELAESRHHIGLQIDSHOGLNMIQOAAIKKNGIVYRSDHCHVH
SFDINASHSDSDGASPDVGSFLYIHMDGEIHHPADGIDGQGGCTPKAKRDLY
SYDIKGAQNIIRNLNDNRSTGRLADRFNMASSMLTQGDGDFKRTARYSPELDNRG
NAAEAFNGTADIYKNIIGAAGETVAGDVOGISECSNIIIVMHGILGSTENKMARIN
DLADANOLKDYAAALIRDAWONPNAOGIEAVSNIFEMAIPIKIGAVRGYGLGGI
TAHPIKRSGMGALALPKGSADVDFNADADAAYAKYPPYHRSNRIRNSIEQYKENTIS
STVPPSRGNKRYLADNRHRTKGYPEFGKGFPPNREKVIKDTKIDIELSGGIPKAP
VPDAPRMEVVDRLNKLITTRDEVKVOELRNQINSNSQHAOLEIRIKLSADIE
NFADMGKFTDSMNDRKAFSRLYSVENGTNPVVEYVEINGRAYIVIRGNRRFAEY
ICRIHEIKFKTKYDFVPNPTSMKPTDVLNDSGNVKKRPRYSK"
/complement(8335. .8799)
/gene="NMB0644"
Query Match 75.7%; Score 1427.8; DB 1; Length 11796;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 1468; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 1 ttggcgatttcccgcaaaataatccctattctgtccatctgcaagtgtgctgcgagt 60
|||||
Db 6506 ttggcgatttcccgcaaaataatccctattctgtccatctgcaagtgtgctgcgagt 6565
QY 61 catgcacgcgcctcagatttggcaacagatccctttatccgcgagttctcgaccgtcag 120
|||||
Db 6566 catgcacgcgcctcagatttggcaacagatcttcttccggcaggttctcgaccgtcag 6625
QY 121 cattcgaacccgcaggaataaccaccttcgcgcagcagggggagcttgcnaagcgc 180
|||||
Db 6626 cattcgaacccgcaggaataaccaccttcgcgcagcagggggagcttgcnaagcgc 6685
QY 181 aacggccatatacgatttggaaacatacaagccatcaagtttggcaccgtatattcaa 240
|||||
Db 6686 agcgcccatatcgatttggaaacatacaagccatcaagtttggcaccgtatattcaa 6745
QY 241 caagcgcgcttgaagaaatalcgcgtacatttgcgcctttccgattccgatacagacaaa 300
|||||
Db 6746 cagggcgccatttaagaaatalcgcgtacatttgcgcctttccgattttccgatacagacaaa 6805
QY 301 ttccattcgcccttcgacaacccatcgcctacattcgcatttcgaagaagccggtagtc 360
|||||
Db 6806 gttcattcgcccttcgacaacccatcgcctacattcgcatttcgaagaagccggtagtc 6865
QY 361 gttgaagcgtatcagccttccatccatcttggagcagatacgaacccatcccgcgac 420
|||||
Db 6866 gttgaagcgtatcagccttccatccatcttggagcagatacgaacccatcccgcgac 6925
QY 421 ggtatgacgagcagcagcagcgtatcccgctccccaagcgcgagagatatatc 480
|||||
Db 6926 ggtatgacgagcagcagcgtatcccgctccccaagcgcgagagatatatc 6985
QY 481 agctacgacataaagcgcttgcaccaaatatccgctcaacctgcagcagacaacgcagc 540
|||||
Db 6986 agctacgacataaagcgcttgcaccaaatatccgctcaacctgcagcagacaacgcagc 7045
QY 541 accggcaacagcgttgcgcgcgtttcccaaatgagcgcgtatcgtgcagcagagagata 600
|||||
Db 7046 accggcaacagcgttgcgcgcgtttcccaaatgagcgcgtatcgtgcagcagagagata 7105
QY 601 ggcgaagcattcaaacgcgcacccgatacagcccgagcttgcagcagatcgcggaatgcc 660
|||||
Db 7106 ggcgaagcattcaaacgcgcacccgatacagcccgagcttgcagcagatcgcggaatgcc 7165
QY 661 gccgaagccttcaaacgcgcactgcagatatacgttcaaaaacatcgcgcgcgcgcagagaa 720
|||||
Db 7166 gccgaagccttcaaacgcgcactgcagatatacgttcaaaaacatcgcgcgcgcgcagagaa 7225
QY 721 atgtgtgcgcgcagcagatgcgtgcagaggtatgaagcgaagcgtcaaatcgtgcgtatg 780
|||||
Db 7226 atgtgtgcgcgcagcagatgcgtgcagaggtatgaagcgaagcgtcaaatcgtgcgtatg 7285
QY 781 caagcgttgtgtgtcttccaccgaacaagaatggcgcatcaacagatttgcagat 840

|||||
Db 7286 CACGGCTTGGTCTGCTTCCACCCGAAAAAAGATGCGCCGATCAACGATTGGCAGAT 7345
OY 841 atggcgcaactcaaaagctatgcccgaacagccatccgagtgattggcagttccaaacccc 900
Db 7346 ATGGCCGAACCTCAAGATATAGCCGACAGCCATCCGGATTGGGCAAGCCAAACCCC 7405
OY 901 aatgcccacaaaggcatagaagccgltcagcaatattcttatgcaagccatcccatcaaa 960
Db 7406 AATGCCCAACAAAGGCATATGAAAGCCGTAGCAATATCTTATATGGACCAATCCCATCAAA 7465
OY 961 gggattggagctgtccggggaataacggttggcgagcatcaaggcaatccctgtcaag 1020
Db 7466 GGGATTGGAGCTGTTCGGGAAAAAATACGACTTGGCGGCATCAGCCACATCTTATCAAG 7525
OY 1021 cggltcgagatggcgcgatcgcatgtccgaagaagaaatccgcgctcagcgcaattt 1080
Db 7526 CGGTGCGAGATGGGCGCATGTGCATGTGCCGAAGGGAATCCGCCCTCAGCGCAATTTT 7585
OY 1081 gccagatgagcatagcccaataaccggtcccttaaccattccgcaataatccgtlcaaac 1140
Db 7586 GCCGATGCGGATACGCCAAATATCCGCTTACATTTCCGAATATTCGCTTCAAC 7645
OY 1141 ttggagagcggttaccgcaaaagaacatcactcctcaacccgtgcgcggtcaaacgyc 1200
Db 7646 TTGGAGAGCGGTTACGGCAAAAGAAACATCACCTCTCAACCGTGCCTGTCMAAGCGC 7705
OY 1201 aaaaatgtcaaacatgagcaaacgacccgaagaacgagtgatccgttttggcgttaaa 1260
Db 7706 AAAAATGTCAAACTGCGAGACCAACGCCACCGAAGACAGGGGTACCGTTTGACGGTAAA 7765
OY 1261 ggglttcgaattttgagaagcagctgtaataatgatacgaagctcgataltcaagaata 1320
Db 7766 GGGTTCGGAATTTTGAGAACACGCTGAATAATGATAGCAACTCGATTTCAAGATTAA 7825
OY 1321 tcgggggagcggtatacctaaggttaagcctgtgtttgatgagaaacgagatggaggtc 1380
Db 7826 TCGGGGGCGGTATACCTTAAGGCTAAGCCTGTGTGATGACCAACCGAGATGGAGGTT 7885
OY 1381 gatagaagaatttaataatgacaactctgtgacaggtggagaanaatttcgagaagc 1440
Db 7886 GATAGCAAGCTTATTAATTAATTAACAACCTGTGAGCAGGTGGAATAAATTTCCAGAAATA 7945
OY 1441 agaagaagaaggtcagagtagtcaagttlaaagccacatgcaagcaagaaatgggaaataa 1500
Db 7946 AGGAACGCTATATTAATAACAGTAACCTTTAGCCCAACATGCTCAACTAGAGAGGGAATTAAT 8005
OY 1501 acaggggttaagtttaatacatttataagttggtgat 1536
Db 8006 AACTTAAATCTGCCGATGAATTAATTTTTCGAGAT 8041

RESULT 4
AX044030 349980 bp DNA linear PAT 24-NOV-2000
LOCUS Sequence 109 from Patent WO006791.
DEFINITION AX044030
ACCESSION AX044030
VERSION AX044030.1 GI:11342914
KEYWORDS
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 349980)
AUTHORS Pizsa,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C.,
Wasmann,V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M.,
Scariato,V., Rappelli,R., Frazer,C.M. and Grandi,G.
TITLE Neisseria genomic sequences and methods of their use
JOURNAL Patent: WO 006791-A 109 09-NOV-2000;
CHIRON CORPORATION (US) : THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
Source location/Qualifiers
1. 349980
/organism="Neisseria meningitidis"

/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces.--seq 1: 1 to
349980 349980 bases--seq 108: 300001 to 649980 349980
bases--seq 109: 600001 to 949980 349980 bases--seq 110:
900001 to 1249980 349980 bases--seq 111: 1200001 to 1549980
349980 bases--seq 112: 1500001 to 1849980 349980 bases--seq
113: 1800001 to 2149980 349980 bases--seq 114: 2100001 to
2272325 172325 bases"
BASE COUNT 82523 a 82940 c 96712 g 87805 t
ORIGIN

Query Match 75.7%; Score 1427.8; DB 6; Length 349980;
Best Local similarity 95.6%; Pred. No. 0;
Matches 1468; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

OY 1 ttgggcatltcccgcaaatatccctattctgtccatctgtagagtgcttcgcatg 60
Db 75614 TTGGGCATTTCCCGCAAAATATCCCTTATCTGTCCATCTGCGCATGCTGCTCCGATG 75673
OY 61 catgcaacgctcagatttggcaaacgataccttattccgcaagttctgacggtcag 120
Db 75674 CATGCACACGCTTCAGATTTGGCAAAAGATTTCTTTATCCGCGCAGGTTCGACCGTCAG 75733
OY 121 catlccgaaccccgacgaggaataaccactatctcgcagcagaggggagcttgcnaagcgc 180
Db 75734 CATTTGCAACCCGACGCGGAATACCACTATTCGGCAGCAGGGGGGAATTCGCCAGCGC 75793
OY 181 aacggccatatcgatgttggaaatacacaagccatcagtttggcaccctgatatcaa 240
Db 75794 AGCGGCATATGCGATTGGGAAAAATACAAAGCCATCATGTTGGCAACCTGATATTTCAA 75853
OY 241 caggcgccgttgaaggaataatcgtctacatgttgcgctttccgataccgggcacaaa 300
Db 75854 CAGCGCGCCATTAAAGGAATATCGCTACATTGTCCGCTTTCCGATCACGGCAGACAA 75913
OY 301 tlccaltcgcccttcgacacacatgctcacaltccgattctgacgaagcaggtagtc 360
Db 75914 GTCCATTCGCCCTTGCACAAACATGCTTCATTCGATTCGTATGAAAGCGGTATGTC 75973
OY 361 gttgaggaattcagccttaccgataccatgttggagatacgaacacatcccgagac 420
Db 75974 GTTGACGGAATTTAGCCTTTACCGCATTCATTTGGAGGATACGAACACATCCGCGCAC 76033
OY 421 ggcatacagggcgcaagggcgcggtctaccgctccccaagcgcgagggatataac 480
Db 76034 GGCATTAGACGGGCCACAGGGCGGGCTATCCGCTTCCAAAGGCGGAGGATATATAC 76093
OY 481 agctacgaataaaaggcgttgcccaaaatalccgctcaacctgacgaacaacgcgagc 540
Db 76094 AGCTACGACATTAAGAGGCGTGGCCCAAAATATCCGCTCAACCTGACCGCAACCGCAGC 76153
OY 541 accggacaacggcttgcgacgcgttttcaacaatggcgcgctatgctgaacgaaggtga 600
Db 76154 ACCGGACAACGGCTTGGCCACGCTTTCACAAATGCGGTATGCTGACGCAAGGAGTA 76213
OY 601 ggcaacgattcaaacgcgcaccccgatacagcccgaagcttgacagatcggaatggcc 660
Db 76214 GCGCAGGATTTAAACGCGCCACCCGATACAGCCCGGAGCTGGACAGATGGGCAATGCC 76273
OY 661 gccgaagccttcaacgycagctgcagatatcgtlcaaaaaacatcaatcggcgcgcagaga 720
Db 76274 GCCGAAGCCTTCAACGGCAGCTGCAGATATCGTTAAAAACATCATCGCGCGGCGAGGAA 76333
OY 721 attgtcgcgcaagcgatgctgtcaggttataagcgaagctcaaacattgtgtcatg 780
Db 76334 ATTGTGCGCGCGCGGAGATGCCCTGCGAGGCATTAAGGGAAGGCTCAAAACATTTGCTGATG 76393
OY 781 caggcgttggtctctcttccacccaacaaagaatgtagcgagatcaacgaatttggcagat 840
Db 76394 CACGGCTTGGGTTGCTCTTCCACCGAAACAGATGCGCGCATCAACGATTTGGCAGAT 76453
OY 841 atgcgcaactcaagaactatgcccgaagcgcataccgcatggtggcagtcacaaacccc 900

|||||
Db 76454 ATGGCCCACTCAAGACATGCGCGAGCAGCCATCCGATTTGGGAGTCCAAACCC 76513
QY 901 aatgcgcacaagcagatagaagccgtcagcaatacttattatgacgacatcccaatcaa 960
Db 76514 AATGCCGCAACAAAGCATAGAACCGCTGACGAATCTTTATGGCAGCATCCCATCAAA 76573
QY 961 gggattggaactctccggggaataatacgccttggcgccatcagcgacatcctgtcaag 1020
Db 76574 GGGATTGGAACTTTGGGGAATAATACGGCTTGGCGCGCATCAGGCACATCTATCAAG 76633
QY 1021 cggctcgcaatggygcgcgcatcgcattgcggaaggaatccgcgcgtcagcgacaattt 1080
Db 76634 CGGTCCAGATGGCGCGGATTCGATTCGCCGAAGGAATCCGCCGTGACGACAAATTTT 76693
QY 1081 gccgatgcgcacatagcgcacaataaccgctcccttaccattcccggaataatcgttcaac 1140
Db 76694 GCGGATGCGGCATACCGCAATACCCGTCCTTACCATTCGCCGAATATCCGTTCAAC 76753
QY 1141 ttggagcaagcttacgagcaagaataacatcacctcctcaaccgtgcgcgcgtcaaacgac 1200
Db 76754 TTGGAGCAGGCTTACGCGCAAGAAACATCAGCTCTCAACCGTGCCTGCAAAACGCC 76813
QY 1201 aaaaatgtcaaacctgycagacccaacgcacccggaagacaagcgctacgcgtttgacgyltaa 1260
Db 76814 AAAAATGTCAAACTGGCAGACCAACGCCACCGCAAGACAGCGTACCGTTGACGATTA 76873
QY 1261 gggcttccgaattttgagaagacacgtgtaataatagatacgaagcgtcgatattcaagaatta 1320
Db 76874 GGGCTTCCGAATTTTGGAAAGCACGTGAATATGATAGAAAGCTCGAATTTCAAAATTA 76933
QY 1321 tcggggggcggtatatacctaagcctaagcctgctgtttgattgagaaacgagattggaagtt 1380
Db 76934 TCGGGGGCGGTTATACCTTAAGGCTTAAGCTGTGTGTTGATGCGAAACCGAGATGGAGTT 76993
QY 1381 gataagaaagcttaataaatctgacaactcgtgagcaaglygagaaatgltcaagaaacg 1440
Db 76994 GATAGGAAGCTTATTAATATGACAACTGCTGAGCAGTGAGAAATAATGTTACAGAAATA 77053
QY 1441 aagaagaagcgctcagctagctggtttaaagccatgagcaagagatggaataataa 1500
Db 77054 AGGAACGGTATATTAACAACTTAACCTTAAGCCACATGCTCAACTAGAGAGGGAATTAAT 77113
QY 1501 acagggctagatttaacttattatagtggtgat 1536
Db 77114 AACTTAATAATCTGCCGATGAATTAATTTTGCAGAT 77149

RESULT 5
A96432 1671 bp DNA linear PAT 07-SEP-2000
LOCUS A96432
DEFINITION Sequence 465 from Patent WO924578.
ACCESSION A96432
VERSION A96432.1 GI:6780108
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1671)
AUTHORS Pizzo,M., Scariato,V., Rappuoli,R., Grandi,G. and Masiagnani,V.
TITLE Neisserial antigens
JOURNAL Patent: WO 924578-A 465 20-MAY-1999;
PIZZA MARIGRADIA (IT); SCARILATO VINCENTO (IT); RAPPUOLI RIMO (IT);
CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)
location/Qualifiers
FEATURES
source 1..1671
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 484 a 440 c 402 g 345 t
ORIGIN
Query Match 65.2%; Score 1231.2; DB 6; Length 1671;

Best Local Similarity 96.3%; Pred. No. 2,7e-281;
Matches 1260; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 1 ttgggacattccgcgaataataatcccttatctgtlccatctatgtaagtgctgtcagatg 60
Db 1 TTGGGATTTCCCGCAAAATATCCCTTATCTGCTCACTGGAGTGTGCTCCGATG 60
QY 61 catgcacagccctcagatttggcaaacgatacccttattccggcaggttctcgaccgtcag 120
Db 61 CATGCACAGCCCTCAAGATTGGCAACGATTCTTTATCCGGCAGGTTCTCGACCGTCAG 120
QY 121 catlccgaacccgcgcggaataataccaactatccgcagcagggggagcttgcacagcgc 180
Db 121 CATTTGAAACCCGACGGAATATACACCTATTGCGCAGCAGAGGGGGAATCTGCGACGCG 180
QY 181 aacgcacataatcgatttgggaataacatacaagccatcaagtttggccaccctgattcaa 240
Db 181 AGCGGTATATCGAATGGGAACATACAAAGCCATATGTTGGCAACTGTTCATCTCAG 240
QY 241 caagcgccgcttgaaggaataatcgcgtacattgtccgctttccgattcaacgagcaaca 300
Db 241 CAGCGCGCCATTAAAGGAATATCGCTACATTTCCGCTTTTCCGATCAGGGCAGCA 300
QY 301 tlccattcgcccttcgacaacacatgcctcaacttcgatttcgagaaagccggtatgcc 360
Db 301 GTCCATTCCCCCTTGACACAACCATGCTCACAATTCGATTCGATTTGTAAGCCGTAATCCC 360
QY 361 gttagaagattcagccttccatccatctattgagagagatacgaataccatcccgagac 420
Db 361 GTTGGACGGATTACGCTTTTACCCATTCATTTGGAGCGGATATGAAACCATCCGCGCAC 420
QY 421 ggtatgagcgagcacaagggcgcggtatcccgctcccaagacgcgagagatataac 480
Db 421 GGTATGACGCGGCACAGGGCGCGGCTATCCGCTCCCAAGCGCGAGGATATATAC 480
QY 481 agctacgacataaagcgcttgcacaaataatccgctcaactgcagcagacaacgcagc 540
Db 481 AGCTACGACATTAAGAGCGCTTGCCCAAAATATCGCCTCAACCTGACGACACCGCAGC 540
QY 541 accggaacaggtctgcgcgcgtttccacaatgcggcgctatgctgagcgaagagata 600
Db 541 ACCGGAACAGGCTGTGCGACCGCTTCCACATACCGGTACTATGCTGACCCAAAGAGATA 600
QY 601 ggcgagatattcaaacgcgcacccgatacagcccgagctgacagatcgcgcaatgcc 660
Db 601 GCGGAGGATTTCAACGCGCAGCTGCGATATGCTCAAAAAACATCATCGCGCGAGAGAA 720
QY 721 attgtgcgcagagcgatgctgcgtgcaaggtataagcgaaagcttaaatgtcgttcat 780
Db 721 ATTGTGCGCGCAGCGGTGCGTGCAGGATTAAGCGAAAGGCTCAAAATTTGTTATG 780
QY 781 caccgcttggtctgcttccaccgaataaagaatgagcgcatcaagatttggcagat 840
Db 781 CACGCGCTTGSGCTGCTTTCACCGAAACAAAGATGGCGGCATCAACGATTTGGCAAT 840
QY 841 atggcgcaactcaaaagactatgcgcagcagcatalccgcgatttggcagtlccaaacccc 900
Db 841 ATGGCGCAACTCAAAAGACTATGCGCGAGACCCATCCGCGATTGGCGAGTCCAAACCCC 900
QY 901 aatgcgcacaagcagatagaagccgtcagcaatacttattatgacgacatcccaatcaa 960
Db 901 AATGCCGCAACAAAGCATAGAACCGCTGACGAATATCTTTACGGAGTATCCCTCAAA 960
QY 961 gggattggaactctccggggaataatacgccttggcgccatcagcgacatcctgtcaag 1020
Db 961 GGGATTGGAACTCTTCCGGGAATAATACGGCTTGGCGCGCATCAGGCACATCTCTCAAG 1020
QY 1021 cggctcgcaatggygcgcgcatcgcattgcggaaggaatccgcgcgtcagcgacaattt 1080
Db 1021 CGGCTCGCAATGgygcgcgcatcgcattgcggaaggaatccgcgcgtcagcgacaattt 1080

```

Db      1021 CGGTCCGAGATGGCGAGATCGCATTTGCCGGAAGGAAATCCGCCCTCAGCGCAATTTT 1080
Qy      1081 ggcgatgcgatacgaacaataaccctcccttaccattccccaataatccgttcaaac 1140
Db      1081 ggcgatgcgatacgaacaataaccctcccttaccattccccaataatccgttcaaac 1140
Qy      1141 ttgagcagcgttacgcgcaaaacacatcacctctctcaacgctgcgcgcgtcaaacgcg 1200
Db      1141 ttgagcagcgttacgcgcaaaacacatcacctctctcaacgctgcgcgcgtcaaacgcg 1200
Qy      1201 aaaaatgtcaaacgcgcgaccaccaacgcaccccgagagcgggtacgcgtttgacgtaaa 1260
Db      1201 aaaaatgtcaaacgcgcgaccaccaacgcaccccgagagcgggtacgcgtttgacgtaaa 1260
Qy      1261 ggggttcgcgaattttgagaaacacgtgaatatgatacgaagctcgata 1309
Db      1261 ggggttcgcgaattttgagaaacacgtgaatatgatacgaagatttaata 1309

RESULT 6
LOCUS   NMA322491 311321 bp DNA linear BCT 04-DEC-2000
DEFINITION
Neisseria meningitidis serogroup A strain Z2491 complete genome;
ACCESSION
AL162754 AL157959
VERSION  AL162754.2 GI:7379424
KEYWORDS
SOURCE   Neisseria meningitidis Z2491.
          Neisseria meningitidis Z2491.
          Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
          Neisseria.
REFERENCE
  1 (bases 1 to 311321)
    Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
    Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
    Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N.,
    Holtoyd,S., Jagels,K., Leather,S., Moule,S., Mungall,K.,
    Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
    Skelton,J., Whitehead,S., Spratt,B.G. and Barrall,B.G.
    Complete DNA sequence of a serogroup A strain of Neisseria
    meningitidis Z2491
    Nature 404 (6777), 502-506 (2000)
2 (bases 1 to 311321)
  Parkhill,J.
  Direct Submission
  Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
  sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
  Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
  Notes:
  Details of N. meningitidis sequencing at the Sanger Centre are
  available on the World Wide Web.
  (URL, http://www.sanger.ac.uk/Projects/N\_meningitidis/).
  Location/Qualifiers
    1. 311321
      /organism="Neisseria meningitidis Z2491"
      /strain="Z2491"
      /db_xref="taxon:122587"
      /note="serogroup: A"
      115..120
      129..587
      /gene="NMA0698"
      129..587
      /gene="NMA0698"
      /gene="NMA0698"
      /note="NMA0698, possible periplasmic protein, len: 152 aa;
      unknown, contains a probable N-terminal signal sequence
      and lies within a region of unusually low GC content"
      /codon_start=1
      /transl_table=11
      /product="putative periplasmic protein"
      /protein_id="CAB83985.1"
      /db_xref="GI:7379425"
      /translation="MTKIKILFLFVFLVLAINLFFRSSDIESGNTQFEYVYD"

FEATURES
  source
    1. 311321
      /organism="Neisseria meningitidis Z2491"
      /strain="Z2491"
      /db_xref="taxon:122587"
      /note="serogroup: A"
      115..120
      129..587
      /gene="NMA0698"
      129..587
      /gene="NMA0698"
      /gene="NMA0698"
      /note="NMA0698, possible periplasmic protein, len: 152 aa;
      unknown, contains a probable N-terminal signal sequence
      and lies within a region of unusually low GC content"
      /codon_start=1
      /transl_table=11
      /product="putative periplasmic protein"
      /protein_id="CAB83985.1"
      /db_xref="GI:7379425"
      /translation="MTKIKILFLFVFLVLAINLFFRSSDIESGNTQFEYVYD"

```

```

misc_feature
  /note="PS00211 ABC transporters family signature"
  /gene="NMA0699"
  /complement(1101..2327)
  /gene="NMA0700"
  /complement(1101..2327)
  /gene="NMA0700"
  /note="NMA0700, possible ribonuclease BN, len: 408 aa;
  similar to SW:RBN.ECOLI (EMBL:J19201), rbn, Escherichia
  coli ribonuclease BN (EC 3.1.-.-), fasta scores: E():
  2.3e-28, 34.3% identity in 268 aa overlap. Longer than rbn
  at the C-terminus. Also similar to TR:085449
  (EMBL:AF067083) Vitreoscilla sp. hypothetical protein (376
  aa), fasta scores: E(): 0.45.0% identity in 333 aa
  overlap. Contains a region similar to NMA0677, fasta
  scores: E(): 4.3e-08, 64.4% identity in 45 aa overlap.
  Contains hydrophobic, possible membrane-spanning regions,
  as does rbn."
  /codon_start=1
  /transl_table=11
  /product="putative ribonuclease BN"
  /protein_id="CAB83987.1"
  /db_xref="GI:7379426"
  /db_xref="SPTREMBL:Q9JVT9"
  /translation="MTFLDRLGLADNKCICAFWYRRFDEERVRQAAASMTFTLL
  ALVPLVTVAVVASIEFVDFDRWSDSEVFNQTIYOGADWVDYINAFEEQNRLLTA
  IGSVDAIYVTSIMLIRTIINTENRIMRVNSORPMWQFLYWMALITFGPLSIGISPM
  VGSVQDAIAGAPQMGALRIATLTETLLIMGLYRPVNPVFPARQAFVAGIATA
  FCLTARSLFTWYMGNPNPGYRSYICAFAPVPELMLNLTMTLVICGAVTSSLSWQ
  GEAFRRGDSKGRFDVYKILLIDRAQKEGKALPVOERRRHINKGIDELGELEKLA
  RHGIYISGRQGWLVLTGADSTELNELFKLYRPLVERDHYNQAVDVMPPCLQTLN
  MTLAEFDAQAKQDQS"
  1376..1385
  /note="Core DNA uptake sequence: gccgctcgaa"
  /label=DUS
  /complement(2417..2426)
  /note="Core DNA uptake sequence: gccgctcgaa"
  /label=DUS
  2444..2453
  /note="Core DNA uptake sequence: gccgctcgaa"
  /label=DUS
  2486..2743
  /gene="NMA0701"
  2486..2743
  /gene="NMA0701"
  /note="NMA0701, possible pseudogene, len: 258 bp; shows
  weak similarity to part of SW:WBA.ECOLI (EMBL:W09166),
  wrba, Escherichia coli trp repressor binding protein (197
  aa), fasta scores: E(): 0.004, 29.3% identity in 82 aa
  overlap. Also similar to part of TR:085450 (EMBL:AF067083)
  Vitreoscilla sp. trp repressor binding protein (fragment)
  (124 aa), fasta scores: E(): 6.6e-14, 56.6% identity in 83
  aa overlap"
  /codon_start=1
  /pseudo

```


Db 143621 GGGATGTGAGCTGTTCCGGGAAAAATACGCGCTTGCGCGCATCAGCGACATCCGTGCAAG 143680
|
QY 1021 CGGTGCGAGATGGGCGGATCGCATTCGCGAAGGAAATCCGCTCGAGCGAATTTT 1080
|
Db 143681 CGGTGCGAGATGGGCGGATCGCATTCGCGAAGGAAATCCGCTCGAGCGAATTTT 143740
|
QY 1081 gccgagtcgagcatatcgcaaaataccgcgtcccttaaccattcccgaaataatccgttcaac 1140
|
Db 143741 GCGGATCGCGCATACGCCAAATACCGCTGCCCTTACCATTCGCCGAATATCCGTTCAAC 143800
|
QY 1141 ttgagagagcgttacgcgcaaaagaataacatcacctccctcaacgcgtgcgcgcgtcaaacgc 1200
|
Db 143801 TTGGAGAGCGGTTACGGGAAAAAACAATCACCTCCTCAACCGTGCCTCGCTCAAAACGGA 143860
|
QY 1201 aaaaatgtcaaacctgcgcagcaaacacccgcgaagcagcggtacccgtttgacggttaa 1260
|
Db 143861 AAGAATGTGAACCTGGCAAAACCAACGCCACCGAAGACCAAAAGTGCCTGTTGACGGTAA 143920
|
QY 1261 ggggttccgaattttgagaagcagctgaataatgatacgaagctcgata 1309
|
Db 143921 GGGTTCCGATTTTGA AAAAGACGTAAATACGATACGAGATTTAATN 143969
|
RESULT 7
AX236465 2304 bp DNA linear PAT 26-SEP-2001
LOCUS Sequence 158 from Patent WO0164922.
DEFINITION AX236465
ACCESSION AX236465.1 GI:15796054
VERSION
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 2304)
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C.C., Maignani,V.C.,
TITLE Heterologous expression of neisserial proteins
JOURNAL Patent: WO 0164922-A 158 07-SEP-2001;
Chilton Spa (IT)
FEATURES
source 1..2304
location/Qualifiers
1..2304
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="961c1-ORF46.1"
BASE COUNT 710 a 619 c 555 g 420 t
ORIGIN
Query Match 63.5%; Score 1198.8; Db 6; Length 2304;
Best Local Similarity 98.1%; Pred. No. 1.3e-273;
Matches 1212; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Db 1369 AGCTTTTACCGCATTCATTGGACGGATACGAACACCATCCCGCGAGCGCTATGACGG 1428
|
QY 433 ccacagagcgcggtatcccgcttcccaaaagcgcgagagatatacagctacgacata 492
|
Db 1429 CCACAGGCGCGGCGGTATCCCGCTCCCAAGGCGCGAGGATATATACAGTACGACATA 1488
|
QY 493 aaagcggttggcccaaaataatccgcttcaacctcgaacgaacacgcgagacacgcgagacag 552
|
Db 1489 AAAGCGTTGGCCCAAAATATCCGCTCAACCTGACCAACCAACCGCACACCGACCGACAAAG 1548
|
QY 553 ctgcgcgacgcttcccaaatgcccgcgcgtcgtatgctgaacgaagagtagcgagcgtatc 612
|
Db 1549 CTTCGCGACCGTTTCCACATGCGGGTAGTATCTGACGCAAGAGTAGGAGGAGGATTC 1608
|
QY 613 aaagcgccaccccgatatacagcccgagctggacagatcgggcaatgcccgcgaagccttc 672
|
Db 1609 AAACGCGCACCCGATATACAGCCCGAGCTGGACAGATCGGGCAATGCGCGGAAGCCTTC 1668
|
QY 673 aacgagctgcagatatacgtaaaacatcatctgcgcgcgagagagaatgttcgcgcga 732
|
Db 1669 AACGCGACTGCGATATCGTTAAACATCATCGGCGGCGAGAGAAATTTGTGGCGCA 1728
|
QY 733 ggcgatgcgctgcaggtataagcgaagcctcaacattctgtcaatgcagcgttggt 792
|
Db 1729 GCGCATGCGCGTGCAGGGCATTAAGCGAAGCCTCAACATTCCTGTCATGCAACGCGTTGGGT 1788
|
QY 793 ctgcttccacccgaaacaagaatggcgcgcacataacgatttgcgagatagcgcaatc 852
|
Db 1789 CTGCTTTCCACCAGAAACAAAGATGGCGCCCATCAACGATTTGGCAGATATGGCCCACTC 1848
|
QY 853 aaagactatgcgcgagcaacatccgcgattgaggagcttcccaaaaccccaatgcgcacaa 912
|
Db 1849 AAAGACTATGCGCGACGACACCTCCCGATTTGGCGCATCCAAAACCCCATATCCGCGCAA 1908
|
QY 913 ggcataagaagccgtcaagaataatcttatgagcagccatcccatcaaaaggtatgagct 972
|
Db 1909 GGCATAGAGAGCGGTCAGCAATATCTTTATGGCAGCCATCCCATCAAAAGGATGTGAGCT 1968
|
QY 973 gtccggggaataatacgcgttggcgagcatcaaggaacatctgtcaagcggtcgcgaatg 1032
|
Db 1969 GTTCGGGAAATATACGGCTTGGCGGCGCATACGCGCATCTCTATCAAGCGGTGCAGATG 2028
|
QY 1033 ggcgcgcatcgcatttgcgcaaaaggaatccgcgcgttaagcgaacaatttgcgcgtgcga 1092
|
Db 2029 GCGCGATCGCATTCGCCGAAGAGGAATCCGCGGTACGAGACAAATTTTCCGATGCGGCA 2088
|
QY 1093 taagccaatatccgctcccttaaccattccggaatatccgltcaaaacttggagcagcgt 1152
|
Db 2089 TACGCCAAATACCCGCTCCCTTACCATTTCCGAATAATCCGTTCAAACTTGGAGCACCGT 2148
|
QY 1153 taagcgaagaataatacaccctcccaacgcgtgcgcgcgttcaaaagcgcaaaatgttcaaa 1212
|
Db 2149 TACGCAAAAGAAACATATACCTCTCAACCGTGCAGCGCTCAACGCGCAAAATGTAAAA 2208
|
QY 1213 ctggcagacccaacgccaacccgaagaacaggcgtacagcttgaacggtlaaaggtttccgaat 1272
|
Db 2209 CTGGCAGACCAACGCCACCCGAAGACAGGCGTACCGTTTGACGCTAAAGGTTTCCGAAT 2268
|
QY 1273 tttagaagcagcgtgaataatgatacgaagctcga 1307
|
Db 2269 TTGAGAAAGCACGTGAATAATGATAGTAACCTCGA 2303
|
RESULT 8
AX239751 2304 bp DNA linear PAT 26-SEP-2001
LOCUS Sequence 53 from Patent WO0164920.
DEFINITION AX239751
ACCESSION AX239751
VERSION AX239751.1 GI:15797386
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.

artificial sequence.
1 (bases 1 to 2304)
AUTHORS Arico,M.B., Comanducci,M., Galeotti,C., Maignani,V., Giuliani,M.M.
and Piza,M.
TITLE Hybrid expression of neisserial proteins
JOURNAL Patent: WO 0164920-A 53 07-SEP-2001;
Chiron Spa (IT)
FEATURES
source location/Qualifiers
1..2304
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="961CL-ORF46.1"
BASE COUNT 710 a 619 c 555 g 420 t
ORIGIN

Query Match 63.5%; Score 1198.8; DB 6; Length 2304;
Best Local Similarity 98.1%; Pred. No. 1.3e-273;
Matches 1212; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 73 tcaatttggaacagatcccttataccggaagttctgacggttcagatcttgagacc 132
DB 1069 TCAATTTGGCAAGATTCCTTTATCCGGCAGGTTCTGACCGCTCAGCATTTGCAACC 1128
QY 133 gagggaaataccaccatctcgacagagggagcttgcacagcgcaagccatc 192
DB 1129 GACGGGAATACCACTATTGGGACAGGGGGAATCTGCGAGCGGCGCATATC 1188
QY 193 ggattggaaacatacaaacgcatcagttggcaccctgataatcaagcgagcgt 252
DB 1189 GGATTGGGAAATATCAAAAGCATCAGTTGGGCAACCTGATTCATACAGGCGGCATTT 1248
QY 253 gaaggaataatcggtatcatatgtccgctttccgatacagggacaaatccattgcgcc 312
DB 1249 AAAGAAATATCGGCTATATGTCGCTTTTCGATCAAGGCGACCAACATCCATTCCTCC 1308
QY 313 ttgcacaacatgccttcacatctcgatctgacagagcggtatgcttcgagatc 372
DB 1309 TTCGACAAACATGCTTCACATTCGATTCGTATGAAGCCGGTATGCCGTGACGGATTT 1368
QY 373 agccttaccgcatcatttggagcgatacgaacacatcccgcgagcggtatgaagg 432
DB 1369 AGCTTTACCGCATCATTTGGAGCAGATACGAACATCCCGCGCATGTGACGGG 1428
QY 433 ccaagagcgagcggtatccgctcccaagcgagagagataatacagcttaacgata 492
DB 1429 CCACAGGCGGCGCTATCCGCTCCCAAGGCGGAGGATATATACAGTACGACATA 1488
QY 493 aaagcgcttgcacaaataatccgctccaactgaacgcaaacgcaagcagcaacgag 552
DB 1489 AAAGCGCTTGCCCAAAATATCCGCTCAACCTGACCGACACACCGGACAAACGG 1548
QY 553 ctgcgacgctttccacaatgcggcgctatgtctgaagcaagagtaaggcgagatc 612
DB 1549 CTTGCCGACCGTTTCCACAATGCCGTACTATGCTGACCCCAAGGACTAGCGAGATTC 1608
QY 613 aaacgcgcaacccgataagcccgagctgagacaagtccggcggaagccttc 672
DB 1609 AAACGCGCACCCGATACGCCCCGAGCTGGACAGATCGGGCATGCGCGGAAGCCCTTC 1668
QY 673 aaagcagctgcagatatcgtcaaaaacacatccgagcgagcaggaagaatctgcgagca 732
DB 1669 AAAGCGCATCGGAGATATGCTTAAAAACATCATTCGGCGGACAGAGAATTTGCGGCA 1728
QY 733 ggcgagtcggtgcaggtataaagcgaagctcaaacattgcttcatgacggtcttgggt 792
DB 1729 GCGGATGCGCTCAGGGCGTAAGCGAAGGCTCAAAATTGCTGTCAATGCGCGCTTGGGT 1788
QY 793 ctgctttcaacgaaaacagaatgagcgcatcaacagatttggcaatatgagcgcaatc 852
DB 1789 CTGCTTTCCACCGAAATACAAAGATGGCGGCATACGATTTGGCAGATATGGCGCAATC 1848
QY 853 aaagactatgccgacagcatalccgagattggcagttccaaaaccccaatgcccagcaaa 912

|||||
DB 1849 AAAGACTATGCCGACAGGCATCCGGATTGGCAGTCCAAAACCCCAATGCCGACAA 1908
QY 913 gcatagaagccgtcgaacataatcttatgcaagccatcccatcaagaaggttggact 972
DB 1909 GGCATAGAACCCGTCACAAATATCTTTATGGACGATCCCATCAAAAGGATTTGGAGCT 1968
QY 973 gtccgggaaatacagcgttggcgacatcagcgacatctctgtcaagcggtcagatg 1032
DB 1969 GTTCGGGGAAATATCGGCTTGCGGCGATCAAGCGACATCTCATTAACCGGTGCGAGATG 2028
QY 1033 ggcgcagctcgcattgcggaaggaatccgcgtcagcgacaaatcttgcgagtcgagca 1092
DB 2029 GCGCGATTCGATTCGGAAGGAAATCCCGCGCACGACAAATTTGCCATCGCGCA 2088
QY 1093 taagcgaataacgctcccttacatcccgaaataatcgttcaacttgaagcagct 1152
DB 2089 TACGCCAAATACCCGTCCTTACCATTCGCCGAATATCCGTTCAAACTTGAACGACGT 2148
QY 1153 taagcgaagaataacatcactcctcaacgctgcccgcgtcaaacgcaaaaatgtcaaa 1212
DB 2149 TACGGCAAAAGAAATACCTCTCTCAACCGTGCAGGCAAAAGGCAAAATGTCAAA 2208
QY 1213 ctggcagaaccaagccacccgaagcagcggtacccgtttgacggtaaaggtttccgaat 1272
DB 2209 CTGCGACAGCAACGCCACCCGAAAGACAGCGGTACCGTTGACGTTAAAGGCTTCCGAA 2268
QY 1273 tttagaagcgcgtgaaatatgataggaagctcga 1307
DB 2269 TTTGAGAAACGACGTGAATATGATGATCACTCA 2303

RESULT 9
AX236421 2028 bp DNA linear PAT 26-SEP-2001
LOCUS AX236421
DEFINITION Sequence 114 from Patent WO0164922.
ACCESSION AX236421
VERSION AX236421.1 GI:15796033
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE
1 (bases 1 to 2028)
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C.C., Maignani,V.C.,
Giuliani,M.M. and Piza,M.C.
TITLE Heterologous expression of neisserial proteins
JOURNAL Patent: WO 0164922-A 114 07-SEP-2001;
Chiron Spa (IT)
FEATURES
source location/Qualifiers
1..2028
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="delta741-ORF46.1"
BASE COUNT 557 a 577 c 531 g 363 t
ORIGIN

Query Match 63.4%; Score 1196; DB 6; Length 2028;
Best Local Similarity 98.3%; Pred. No. 6.1e-273;
Matches 1208; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 71 cctcagatttgcgaacgatacccttataccggaagttctcgaacgctcagatctcgaaac 130
DB 773 COTCAGATTGGCAAAACGATTTCTTTATCCGGCAGAGTCTCGACGCTCATTTGGAAC 832
QY 131 ccgacgggaataatacactatctcgacagcagggggaagcttgcacagcgcaagcgcata 190
DB 833 CCGACGGGAAATACCACTATTTCGACAGCGGGGAACCTTGCCGAGCGACGCGCATTA 892
QY 191 tcgatttggaaataatacaaacgcatcagttggccacactgatatcaacagcgcgcg 250
DB 893 TCGATTGGGAAATATCAAAAGCATCACTGTGGCAACTGATGATTCAAACAGCGGCCA 952

QY	251	tgaaggaataatcgcgtacatgtgtccgcttllcogalcaagggcaacaatttcattcgc	310
Db	953	TTAAAGGAATATTCGGCTACATTTGTCGGCTTTCCGATCCACGGGCACGAATTCATTCC	1012
QY	311	ccctcgacaacacagctccatcattccgattctcgaaagccggttagtccgtttagcgtat	370
Db	1013	CGTTGCAACCATGCTTCACATTCCGATTTCTGTGAAGCCGGTAGTCCTGTAACGAT	1072
QY	371	tcaagcttlaacgcgcatccatttggacgagatagacacacatcccgccgacgagctatgcg	430
Db	1073	TTTAGCCTTTACCCCATTCATTGGGACGGATACGAACACCATCCCGACGGCTATGACG	1132
QY	431	ggccacagggccggcggtatcccgctcccaagcgcgagggatataatacagctagaca	490
Db	1133	GGCCACAGGGCGGGGCTATCCGCTCCCAAAAGGCGGAGGATATATACGTACGACA	1192
QY	491	taaaaggcgtttgcaccaataatccgctcaacctgacgcgaacaacccgacacccgacac	550
Db	1193	TAAAGGCGTGTGGCCAAATATTCGGCTCAACCTGACCGACGACACCGGACGACAC	1252
QY	551	ggcttgcgcgacgctttccaatggccggcgctatgtctgacgcaagsgtaggcgacgcat	610
Db	1253	GGCTTGCCGACCGCTTTCCCAATGGCCGGTGTGTAGCTGACGCAAGGAGATGGGCGACGAT	1312
QY	611	tcaacgcgcgcaccccgatatacagcccgcagcttgcacagatccgggaatggccgcgaagct	670
Db	1313	TCAACGCGCCACCCGATACAGCCCGGAGCTGACAGATGGGCAATGCCCGCAAGCCT	1372
QY	671	tcaacggcaetgagatatactcgtcaaaaaatcatcaggcgcggaagaaattgtcggcg	730
Db	1373	TCAACGGCACTGGAGATATGTTAAAAACATCATCCGGCGCGGAGGAAATTTGCGCG	1432
QY	731	caagcgcatgcccgcgcaggggtataagcgaagggccaaacattgtctgtaatacagctctg	790
Db	1433	CAGGCGATGCCGTGCAAGGGCATTAAGCGAAGCTCAAAACATTGCTGTATGCAAGCGCTTG	1492
QY	791	gtctgcttccacgcgaaaaaagatggcgcgcatccaagatttgcagatatgtgcgcaac	850
Db	1493	GTCGTGTTTCCACCGAAACCAATGCGCGCATCAACGATTTGGCGAGTATGCGCGAAC	1552
QY	851	tcaaaacatattgcggaagcagcatccgcattggcgagctccaaccccaatggcgcac	910
Db	1553	TCAAAAGACTATGCCGACGACGCCATCCGGATTGGCGAGTCCAAACCCCAATGCCGAC	1612
QY	911	aaggaatagaagcgtccaagcaatatttlatggcagccatcccatcaaaagggataggag	970
Db	1613	AAGGCTATGAAGGCGGTGAGCAATATCTTTATGGCAGCGCATCCCATCAAAAGGAGTTGAG	1672
QY	971	cgtgcggggaaaaatacgcgcttggcgcgcatcaaggcaatccctgtccaagcgttcgcaga	1030
Db	1673	CGGTTCGGGAAAAATACGGCTTGGGGGATCAACGGCACATCTTATCAACGGGCTCGAGA	1732
QY	1031	tgggcgcagctgcattgcggaagggaaatccgcgttcaagcgaacaatttgcgcgattggg	1090
Db	1733	TGGGGCGGATTCGATTGCCGMAAGGAAATCCGCCGTCAACGCAAAATTTTGGCGATGGG	1792
QY	1091	catacgcaaataccgcgtcccttaccattccccaataataccgtltcaaaacttggagcagc	1150
Db	1793	CATAGCGCAAAATACCGTCCCTTACCATTCGGAATATCCGTCAAACTTGAGAGAGC	1852
QY	1151	gttaccgaaagaaacatcacctctcttaaccgltgcgcgcgttcaaacggcaaaaatgtca	1210
Db	1853	GTTTACGGCAAAATAATCATCTCTCAACCGTGGCGCGTCAACGGCAAAAATGTCA	1912
QY	1211	aactggcagaccaacgcaccccggaagacggcgtaccgltttagcgttgaaggttttcga	1270
Db	1913	AACGTGCAACCAAGCCACCCGAAACAGAGCGGTACCGTTTGAACGTTAAAGGTTTCGA	1972
QY	1271	atttgagaagcagtgtaatatgtatgcg	1299
Db	1973	ATTTTGAGAAACGCTGAATATATGATACG	2001

[illegible]

QY 731 caagcgatgcgctgcaaggatataagcgaagcctcaaaacattgctgtcaatgcaagcgttgg 790
|||||
Db 1433 CAGGCGATGCCGTCAGAGGCGATTAAGCGAGGCTCAAAACATTGCTGTCATGCAAGCGCTTGG 1492
QY 791 gtcctcttcacacgaaataaagaatgagcgacatcaacgatttggcaataatggcgcaac 850
|||||
Db 1493 GTCGTCTTCCACGGAATAAGATGGCGGCATCAACGATTGTGGACATATTGGCGCAAC 1552
QY 851 tcaaaagactatgcgcgaagcagccatccgcgcatgaggcagtcacaaaccccaatgcgcac 910
|||||
Db 1553 TCAAAAGACTATGGCGGCGAGCGCATTCGCGATTGGCGATTCGCAAAACCCCAATGGCGGAC 1612
QY 911 aagcctagaagccgctgaagaataatttatggaagccatcccaatcaaaaggatctgag 970
|||||
Db 1613 AAGGCAATGAAGCGCGTACGACATATCTTTATGCGAGCCATCCCATCAAAAGGATTTGGAG 1672
QY 971 ctgtccgggaaaaataacgcttggcgcgacatcaacgacatctgttcaagcggttcgcaga 1030
|||||
Db 1673 CTGTTGGGGAATAATAGGGCTTGGCGGCATCAGCGCACATCTTCAAGCGGTGCGAGA 1732
QY 1031 tggcgcgatcgcatgycgcgaaggaaatccgcgcgtcagcgacaattttgcgcatgcyg 1090
|||||
Db 1733 TGGCGCGATCGCATTCGCGAAGGAAATCCGCGTCAGGACAAATTTTGGCATGGCG 1792
QY 1091 catacgcaaataccgctcccttaccatccacccgaataatccgttcaacttggagcagc 1150
|||||
Db 1793 CATACGCCAAATACCGCTCCCTTACCATTTCCGAAATATCGTTCAAACTTGGAGACAC 1852
QY 1151 gttaacgcaaaagaaacatcacctctcaacgctgacgctgcaaaacgcaaaatgtca 1210
|||||
Db 1853 GTTACGGCAAAAGAAACATCATCTCTCAACCGTCCGCGCAAAAGCAAAATGTCA 1912
QY 1211 aactggaagcaaacgaccccggaagaagcgatcagtttgcaggttaaggtttcgca 1270
|||||
Db 1913 AACTGGAGACCAACGCCACCGGAGACAGCGGTACCGTTTACGCTAAAGGTTTCCGA 1972
QY 1271 atttgagaagcacgtgaataatgatacg 1299
|||||
Db 1973 ATTTTGAAACACGCTGAATATGATACG 2001
RESULT 11
AX236407 4425 bp DNA linear PAT 26-SEP-2001
LOCUS AX236407
DEFINITION Sequence 100 from Patent WO0164922.
ACCESSION AX236407
VERSION AX236407.1 GI:15796026
KEYWORDS
ORGANISM
SOURCE
synthetic construct.
synthetic construct.
artificial sequence.
REFERENCE
1 (bases 1 to 4425)
AUTHORS
Atico,M.B., Comanducci,M.C., Galeotti,C.C., Masignani,V.C.,
Guillani,M.M. and Pizzi,M.C.
TITLE
Heterologous expression of neisserial proteins
JOURNAL
Patent: WO 0164922-A 100 07-SEP-2001;
Chiron Spa (IT)
FEATURES
source
1..4425
location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="deltaG983-ORF46.1"
BASE COUNT 1165 a 1259 c 1219 g 782 t
ORIGIN
Query Match 63.4%; Score 1196; DB 6; Length 4425;
Best Local Similarity 98.3%; Pred. No. 6,8e-273;
Matches 1208; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 71 cctcagattggcaaacgcttcccttaccggaagtttcgacgcgttcaacatttgaac 130
|||||
Db 3170 CCTCAGATTGGCAACGATTTCTTTATCCGCGAGGTTCTCGACGTCAGCATTTTCGAC 3229

QY 131 ccgacgggaataataccactatcgcgacgaggggagcttgcnaagcgcaacggccata 190
|||||
Db 3230 CCGACGGGAATAATCCACACTTTTGGCGAGCAGGGGGGAACCTTGCGAGCGGCGCATTA 3289
QY 191 tccgatttggaaacatacaaaagccatcagtttggccaccttgaatcaacaagcgccg 250
|||||
Db 3290 TCGGATTGGGAAAAATACAAAGCCATCAGTTGGGCAACCTTGATATTCMAAGCGCGCCA 3349
QY 251 ttgaaggaataatccgctacatggttcgcgtttccgatacgaaggacaaattccattcgc 310
|||||
Db 3350 TTAAAGGAATATCGGCTACATTGTCGCTTTTCGATTCAGGCGACGAAGTCCATTCCC 3409
QY 311 ccttcgacaacatgctctcaatccgatttcgaatcgaagccggtagttccgttgaagat 370
|||||
Db 3410 CTTTCCGACAAACATGCTTCATCTCCGATTTCTGATGAAGCGGATAGTCCGTTAGCGAT 3469
QY 371 tcagccttaccgcatccatcttggagcagatacgaacacatcccgccgacgactaagc 430
|||||
Db 3470 TTAGCTTTACCGCATTCATTTGGAGCGGATACGAACCATCCGCGCATATGACG 3529
QY 431 ggcacagggcgcggtatccgctccaaaggcggaaggatataacagctacgaca 490
|||||
Db 3530 GGCACAGGGCGGCGCTATCCGCTCCAAAGGCGGAGGATATATACACTACGACA 3589
QY 491 taaagcgcttgcacaaataatccgctcaacttgaacggaacacgcaacgcaac 550
|||||
Db 3590 TMAAGCGCTTGGCCCAAAATATCCGCGTCAACCTGACCGCAACCGCGACGACAAAC 3649
QY 551 ggccttcgacacgcttcccaaatgacgagcgtatgctgagcgaaggatlaagcgacgat 610
|||||
Db 3650 GCGTTGGCGACCGTTTCCACATCCGCTAGTACTACGCAAGGATATAGCGACGAGT 3709
QY 611 tcaaacgcygcaccccgatatacagcccgagctggaacagatcgagcaatgcgcgaagcct 670
|||||
Db 3710 TCAACGCGCCACCCGATACACCCCGAGCTGACAGATCCGGGCAATGCCGGAAGCCT 3769
QY 671 tcaacgcaactgacagataatgcaaaacatcatcgcgcgagcgaaggaatatttcgcgcg 730
|||||
Db 3770 TCAACGGCAGCTGAGATATCGTTAAAAACATATCGGCGCGCAGAGAAATTTGCGGGG 3829
QY 731 caagcgatgcgctgcaaggtatgaagcggaagctcaaaactgtctgacgagcttgg 790
|||||
Db 3830 CAGCGATGCCGCTGACAGGCGATTAAGCGAAGGCTCAAACTTCTGTATGACGCGCTTGG 3889
QY 791 gtcgtcttccacgaaacaaagatgagcgcatcaacgatttggcagatattgycgcaac 850
|||||
Db 3890 GTCGTCTTCCACCGAAACAAAGATGGCGGCATCAACGATTGGCAGATATGGCGCAAC 3949
QY 851 tcaaaagactatgcgcgaagcagccatccgcgcatgaggcagtcacaaaccccaatgcgcac 910
|||||
Db 3950 TCAAAAGACTATGGCGCACACGACCATCCGATTTGGCGAGTCCAAAACCCCAATGGCGGAC 4009
QY 911 aagcctagaagccgctggaagaatatttlatgacgacatcccatcaaaaggattggag 970
|||||
Db 4010 AAGGCAATGAAGCGGTGACGACATATCTTTATGAGCAACCATTCGCCATCAAAAGGATTGGAG 4069
QY 971 ctgtccgggaaaaataacgcttggcgcgcatcaacgacacatccgttcaagcggttcgaga 1030
|||||
Db 4070 CTGTTGGGGAATAATAGGGCTTGGCGGCATCAGCGCACATCTTATCAAGCGGTCCGACA 4129
QY 1031 tggcgcgatcgcatgycgcgaaggaaatccgcgcgtcagcgagaaatatttgcgcatgcyg 1090
|||||
Db 4130 TGGGCGCATCGCATTTGCGAAGGGAATTCGCGTCAGCACAATTTTGGATGGAGCGG 4189
QY 1091 catacgcaaataccgctcccttaccatcccgaaataatccgttcaaaacttggagcagc 1150
|||||
Db 4190 CATACGCCAAATACCGCTCCCTTACCATTCGGAATATTCGTTCAAACTTGGAGCAGC 4249
QY 1151 gttaacgcaaaagaaacatcacctctcaacgctgacgctgcgcgttcaaacgcaaaatgtca 1210
|||||
Db 4250 GTTACGGCAAAAGAAACATCACTCTCTCAACGCTGCGCGCTCAAAACGCGCAAAATGTCA 4309

QY 1211 aactgagacacacgcccacgaagacagcgttacccgtttgacggttaagagtttcga 1270
|||||
Db 4310 AACTGAGACAGCAACGCCACCGAAGACAGGGGTACCGTTTACGGTAAAGGTTTCCGA 4369
QY 1271 atttggagaacagtgtaatatgatacg 1299
|||||
Db 4370 ATTTTGAGAAACGCTGAATATGATACG 4398

RESULT 12
AX239715
LOCUS AX239715 4425 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 17 from Patent WO0164920.
ACCESSION AX239715
VERSION AX239715.1 GI:15797369
KEYWORDS
SOURCE synthetic construct.
ORGANISM
REFERENCE 1 (bases 1 to 4425)
AUTHORS Arico,M.B., Commanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.
and Piza,M.
TITLE Hybrid expression of neiserial proteins
JOURNAL Patent: WO 0164920-A 17 07-SEP-2001;
Chiron Spa (IT)
FEATURES
source 1..4425
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="delta983-ORF46.1"
BASE COUNT 1165 a 1259 c 1219 g 782 t
ORIGIN

Query Match 63.4%; Score 1196; DB 6; Length 4425;
Best Local Similarity 98.3%; Pred. No. 6.8e-273;
Matches 1208; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 71 cctcagatttgcgaacagatcccttataccgacaggttcctgacccgttaagcatttcgaac 130
|||||
Db 3170 CCMCAGATTGTGGCAAAAGATTTCTTTATCCGGCGAGTTCTCGACCGTACGATTTCCGAC 3229
QY 131 ccgacgggaataatccacatcttcgacgaagggggagcttgcnaagcgcaagggcaca 190
|||||
Db 3230 CCGACGGGAAATACCAACGATTCGCGAGCGGGGAACTTGCCGACGCGCATTA 3289
QY 191 tcgattgggaaacatacaaacgcatcagttggccacatgatattcaaaagcgcgccg 250
|||||
Db 3290 TCGGATTGGGAAAAATACAAAGCCATCATGTTGGGCAACCTGATGATTCAACAGGCGGCCA 3349
QY 251 ttgaaggaatatcggctacatgttcgcttcccgatcacgagggcacaattccatccgc 310
|||||
Db 3350 TTAAAGGAATATCGGCTACATGTCCGCTTTCCGATCACGGGACAGAGTCCATTCC 3409
QY 311 ccttcgacacacatgctacatcttcgattctgaagagccggtatgtccggttgaagcat 370
|||||
Db 3410 CTTTCGACACACATGCTCATTCGATTCTGATGAAGCGGATGACCCGTTGACCGAT 3469
QY 371 tcaagccttacgcatacttggagcagatacaacaaccctccgcccagcgctatagc 430
|||||
Db 3470 TTAAAGCTTACCGCATTCATTTGGAGAGATACGACACACCTCCCGGACGCTATACG 3529
QY 431 ggcacacagggcgcgctatcccgctcccaagggcgagagatatatacagctacgaca 490
|||||
Db 3530 GGCACAGAGCGCGGCTATCCCGCTCCCAAAGCGCGAGGATATATACAGCTACGACA 3589
QY 491 taaagcgcttgcccaaaatatccgctcaacctgacgcgaacacgcgacagcggaaac 550
|||||
Db 3590 TAAAGGCGCTTCCCAAAATATCCGCTCAACCTGACGACAAACCGACGCGGACAC 3649
QY 551 ggcttcgcgacgctttcccaaatgcgcgcgtatgtgaacgaagagtagtgaagcgat 610
|||||
Db 3650 GGCTTGCCGACCGTTTCCACATGCGCGGTAGTATGCTGACGCAAGGATGAGGCGACGAT 3709

QY 611 tcaagcgcccaaccgatacaagcccagcttggaacagatcgggaatgcgcgcgaagcct 670
|||||
Db 3710 TCAACCGCGCCACCCTGGATACAGCCCCAGCTGAGACAGATCGGGCAATCGCCCGAAGCCT 3769
QY 671 tcaagcgacatgcagatatcgtcaaaacatcatcggcgcgaggggaagtatttcgagc 730
|||||
Db 3770 TCAACGGCACTGCAGATATCTTTAAATAATCATCTGGGGGCGGAGAAATTTGCGGGC 3829
QY 731 caagcgatgcgctgcgaaggtataagcgaaggtcaaaacattgtctcaagcgcttgc 790
|||||
Db 3830 CAGGCGATGCGCGTGCAGAGCATTAAGCAGAGCTCAAAACATTGCTCATGACGCGCTTGG 3889
QY 791 gctgtcttcaccgcgaagaagaatgycgcgcacatcagatttggcagatataggcgcaac 850
|||||
Db 3890 GTCTGCTTTCCACCAGAAACAGATGGCGGCATCAACGATTTGCGAGATATGCGCGCAGC 3949
QY 851 tcaagactatgcgcgcgcgcgcacatccgcgatttggagagtcocaaaccccaatgcgcac 910
|||||
Db 3950 TCAAAAGACTATGCGCGACGACGACCATTCGCGATTGGGACATCCCATCAAGGAGATTGGAG 4009
QY 911 aaggaataagaagcgttcgaacatacttataggcagccatcccatcaaaaggtattgag 970
|||||
Db 4010 AAGGCAATAGAGCGCTGACGACGACCATTCGCGATTGGGACATCCCATCAAGGAGATTGGAG 4069
QY 971 cgttcgggggaataatcgtgttggcgacatcaagcacaatcctgttaagcgctgcaga 1030
|||||
Db 4070 CTGTTGGGGAATAATACGGCTTGGCGGCAATCCGACATCTTATACCGGCTGCGAGA 4129
QY 1031 tggcgcgatcgcgcatgcgcgaaggggaatccgcgcgtcaagcacaatttggcagatgcg 1090
|||||
Db 4130 TGGGCGCGATTCGATTGCCGAAGGGAATCCGCCGTACAGCAATTTTCCGATGCGG 4189
QY 1091 catagcacaataccgctcccttaccatcccgaaatatacgttcaaatgtgaagcagc 1150
|||||
Db 4190 CATACGCCCAAAATACCGCTCCCTTACCATTTCCGAAATATCCGTTCAACTTGGAGCAGC 4249
QY 1151 gttacggcaagaataatcaactcctcaacgctgcgcgcgttcaaacgcaaaatgtca 1210
|||||
Db 4250 GTTACGCGCAAGAAACATCACTCTCTCAACCGTCCGCGCTCAAAACGGCAAAATGTCA 4309
QY 1211 aactgacagaccacgcaccccgaaagacagcgcttacggttgcaggtlaagggtttcga 1270
|||||
Db 4310 AACTGACAGACCAACGCCACCGGAAGACAGGGCTTACCGTTTGAAGGAGGTTTCCGA 4369
QY 1271 atttggagaacagcgtgaatatgatacg 1299
|||||
Db 4370 ATTTTGAGAAACGCTGAATATGATACG 4398

RESULT 13
AX236447
LOCUS AX236447 2019 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 140 from Patent WO0164922.
ACCESSION AX236447
VERSION AX236447.1 GI:15796045
KEYWORDS
SOURCE synthetic construct.
ORGANISM
REFERENCE 1 (bases 1 to 2019)
AUTHORS Arico,M.B., Commanducci,M.C., Galeotti,C., Masignani,V.C.,
Giuliani,M.M., and Piza,M.C.
TITLE Heterologous expression of neiserial proteins
JOURNAL Patent: WO 0164922-A 140 07-SEP-2001;
Chiron Spa (IT)
FEATURES
source 1..2019
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="ORF46.1-741"
BASE COUNT 555 a 571 c 531 g 362 t
ORIGIN

[illegible]

QY	1093	tacggcaaatcccgctcccttaccattcccgaaataatcgtttcaacttggagcagcgt	1152
Db	1024	TACGCCAAATACCCGTCCTCCCTTACCATTCCTCCGAATATTCGTTAAACTTGGACAGCGT	1083
QY	1153	tacggcaaaaataataactactccctcaacccgttgcgcgcgttaaacgcgaataatgtcaaa	1212
Db	1084	TACGGCAAAAGAAATCATCACTCTCTCAACCTGCGCCGCTCAAAAGCGCAAAAATGTCAA	1143
QY	1213	ctggcagacccaacgcaccaccggaagacagcgcttacggtttgacggtgaagggtttccgaat	1272
Db	1144	CTGGCAGACCAACGCGCACCCCGAAGACAGCGCGTACCGTTGACGGTAAAGGGGTTCCGCAAT	1203
QY	1273	tttgagaagcagcgttgaatatgatcgc	1299
Db	1204	TTTGAGAAGCACGCTGAATATGATGACG	1230
RESULT	14		
LOCUS	AX239733	2019 bp	DNA linear PAT 26-SEP-2001
DEFINITION	AX239733	Sequence 35 from Patent WO0164920.	
ACCESSION	AX239733		
VERSION	AX239733.1	GI:15797377	
KEYWORDS			
SOURCE		synthetic construct.	
ORGANISM		synthetic construct.	
REFERENCE		artificial sequence.	
AUTHORS		1 (bases 1 to 2019)	
		Arico,M.B., Comanducci,M., Galeotti,C., Maignani,V., Giuliani,M.M.	
		and Pizzi,M.	
		Hybrid expression of neisserial proteins	
		Patent: WO 0164920-A 35 07-SEP-2001;	
TIME			
JOURNAL		Chiron Spa (IT)	
FEATURES			
source		1..2019	
		/organism="synthetic construct"	
		/db_xref="taxon:32630"	
		/note="ORF46.1-741"	
BASE COUNT	555 a	571 c	531 g 362 t
ORIGIN			
Query Match	63.3%;	Score 1194;	DB 6; Length 2019;
Best Local Similarity	98.3%;	Pred. No. 1.8e-272;	
Matches 1206; Conservative	0;	Mismatches 21;	Indels 0; Gaps 0;
QY	73	tccagatttggcaaaagcatcccttataatcggcgcaggttctcgaccgtcagcatttcgaacc	132
Db	4	TCAGATTTGGCAAAACGATCTTTATCCCGCAGGTTCTCGACCGTCAGCATTTGCAACC	63
QY	133	gaagggaaatataccactatttcggcagcagggggagcttgcacagcgcaagcgaatc	192
Db	64	GACGGGAAATACCACTATTTCGGCAGCAGGGGGAACTTGCCGAGCGCAGCGCCATTATC	123
QY	193	ggaattggaaacaataaagccatacagtttgggcacacttgatgttcaacagggcgcgtt	252
Db	124	GGATTGGGAAATATACAAAGCCATCATGTTGGGCAACCTGATGATTCACACAGCGGCGCAT	183
QY	253	gaaggaataatcggttacctatgttcgcgttttccgatcaaggggcacaatltccattcgccc	312
Db	184	AAAGGAAATATCCGCTACATATTGCCCTTTTCGATTCACGGCAGCAAGTCATTCCCCC	243
QY	313	tttcgacaacataatccataatccatcgcatttcgacgaagccggtatgctccglttgacgattc	372
Db	244	TTTCGACACCAATGCTCTACATTCCTCGATTCTGATGAAACCGCGTATGTCCTCGTTGACGATT	303
QY	373	agcctttaccgcataatcatltygaacgatacgaacacacatcccgccgacggtctatgaagg	432
Db	304	AGCCTTACCGCATTCATTTGGACGGATGATGACACACCATCCCGCAGCGCTATGACGGG	363
QY	433	ccacaaggcggtgctatcccgcttccaaaggcgcgaggtatataatcagctatcgacata	492

Db 364 CCACAGGGGGGGCTATCCGCTCCAAAGCGCGAGGGATATATACAGTACACATATA 423
QY 493 aaagcgcttgcccaaatatccgcctcaacctgaccgcaaacgcaacgcaacgcaacg 552
Db 424 AAAGCGCTTGCCCAAAATATCCGCTCAACCTGACCGACACCGACGACGACGACG 483
QY 553 ctgacgacgcttcccaaatgacgcaacgcaacgcaacgcaacgcaacgcaacg 612
Db 484 CTTGCCACCGCTTCCAAATGCGGCTAGTATGCTGACGCAAGAGTATGCGACGATTC 543
QY 613 aaacgcaaccccgatcacgcccgaagctgacacgacgcaacgcaacgcaacgcttc 672
Db 544 AAAGCGCCACCGCATACGCGCGGAGCTGAGACAGATGCGGCAATGCGCGCAAGCTTC 603
QY 673 aacgcaactgacagatcgtcaaaaacacacacacacacacacacacacacacacac 732
Db 604 AACGCACTGCAGATATCGTTAAAAACATCATCGCGCGCGGCAAGAAATTCGCGCA 663
QY 733 ggcgacgctgacgaggtatagcgaagctcaaacatgctgcatgcaacgcttgagtt 792
Db 664 GCGGATGCCGTGCGAGGCGATAGCGCAAGGCTCAAAACATTCGTCATGACGCGTTGGGT 723
QY 793 ctgcttccacccgaacaaagaatgacgcaacacacacacacacacacacacacacac 852
Db 724 CTGCTTCCACCGCAAAACAGATGGCGGCAATCAAGATTTGGCGAGATATGGCGCACTC 783
QY 853 aaagactatgcccgaagcagcaltccgcatgtggcagctccaaaccccaatgcccacaa 912
Db 784 AAAGACTATGCGCGAGCGACCATCGCGGATTTGGCGAGTCCAAACCCCAATTCGCCACAA 843
QY 913 ggcatagaagccgctcagcaatattctatgacgacacatcccatcaaaaggatggagct 972
Db 844 GGCATGGAACCGCTCACCAATATCTTTATGCGACCAATCCCATCAAAAGGATTTGAGACT 903
QY 973 gtcgagggaaaataacgcttgagcgacacacacacacacacacacacacacacacac 1032
Db 904 GTTCGGGGAAATATACGCTTGCGGCGCATCACGACATCTTATCAACGGCTCGAGATG 963
QY 1033 ggcgacatcgcatgtccgaagaagaaatccgcgctcagcgacaaattgtccgacgca 1092
Db 964 GGGCGCATCGCATTTGGCGAAAGGGAATTCGCGCTCAGCGCAATTTTGGCGATGCGGCA 1023
QY 1093 tacgcaaataccgtcccttaccatcccgcaatattccgttcaaaccttgagagcgt 1152
Db 1024 TACGCCAAATACCGCTGCTTACCATTTCCGAAATATCCGTTCAACTTGACAGCGCT 1083
QY 1153 tacgcaagaagaacatcacctcctcaacgctgacgcaacgcaacgcaacgcaacgca 1212
Db 1084 TACGCAAGAAACATCATCTCTCAACGCTGCGCGCTCAACGGCAAAATGTCAAA 1143
QY 1213 ctgacgac 1272
Db 1144 CTGGCGACCAACGCGACCGGAGACAGGCGTACCGTTGACGCTAAAGGCTTCCGAAAT 1203
QY 1273 ttggaagaacgctgaataataatacag 1299
Db 1204 TTTGAGAAACAGCTGAATATGATACG 1230

RESULT 15
AX236451 2256 bp DNA linear PAT 26-SEP-2001
LOCUS AX236451
DEFINITION Sequence 144 from Patent WO0164922.
ACCESSION AX236451
VERSION AX236451.1 GI:15796047
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE
1 (bases 1 to 2256)
Arico,M.B., Comanducci,M.C., Galeotti,C.C., Maignani,V.C.,
Guilliani,M.M. and Pizsa,M.C.,
Heterologous expression of neisserial proteins

JOURNAL Patent: WO 0164922-A 144 07-SEP-2001;
Chiron Spa (IT)
FEATURES
source location/Qualifiers
1..2256
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="ORF46.1-961c"
BASE COUNT 698 a 608 c 545 g 405 t
ORIGIN

Query Match 63.3%; Score 1194; DB 6; Length 2256;
Best Local Similarity 98.3%; Pred. No. 1,9e-272;
Matches 1206; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 73 tcgatttgcaaacagatcccttaccgcaagtttcgaacgctcagacattcgaaacc 132
Db 4 TCAGATTTGGCAAAACGATCTTTATCCGGCAGGTTCTCGACCGTCAAGCATTTGAACCC 63
QY 133 gacgggaataacacactatccgacgagggagggttgcmaagcgaacgacatc 192
Db 64 GACGGAAATACCACTATTGCGAGCGGGAACCTTCCGCGACGCGCATATC 123
QY 193 ggaattggaacacacaaagcaccatgctggcaccatgaltcaacagcgccgt 252
Db 124 GATTTGGGAAAAATACAAAGCCATGATTGGCAACTGATGATTCAAACAGCGCCCAT 183
QY 253 gaaggaataatcgctacattgtccgctttccgacacagggcaacaaatccatccgc 312
Db 184 AAAGAAATATCGGTACATTTGTCCTTTCCGATCAGCGGCAAGATTCATTCCTCC 243
QY 313 tccgaacacacacgcccacacatccgatttcgaagaagcgggtgcttcgacgagtc 372
Db 244 TTGCAACACCATGCTCATATTCGATTTCTGATGAACCGGCTAGTCCGTTGAGGGATT 303
QY 373 agccttaccgacatccattggaagacacacacacacacacacacacacacacacac 432
Db 304 AGCTTTACCGCATCCATTTGGGAGGATGACGACACCATCCCGCGGCTATGACGGG 363
QY 433 ccacagggggggggtatccgctcccaaaagggagggatataacagctacacata 492
Db 364 CCACAGGGGGGGGCTATCCGCTCCAAAGGCGGCAAGGATATATACACTACGACATATA 423
QY 493 aaagcgcttgcccaaatatccgcctcaacctgaccgcaaacgcaacgcaacgcaacg 552
Db 424 AAAGCGCTTGCCCAAAATATCCGCTCAACCTGACCGACACCGGACGACGACGACG 483
QY 553 ctgacgacgcttcccaaatgacgcaacgcaacgcaacgcaacgcaacgcaacgcaacg 612
Db 484 CTTGCCACCGCTTCCAAATGCGGCTAGTATGCTGACGCAAGAGTATGCGACGATTC 543
QY 613 aaacgcaaccccgatcacgcccgaagctgacacgacgcaacgcaacgcaacgcttc 672
Db 544 AAAGCGCCACCGCATACGCGCGGAGCTGAGACAGATGCGGCAATTCGCGCAAGCTTC 603
QY 673 aacgcaactgacagatcgtcaaaaacacacacacacacacacacacacacacacac 732
Db 604 AACGCACTGCAGATATCGTTAAAAACATCATCGCGCGGCAAGAAATTTGCGCGCA 663
QY 733 ggcgacgctgacgaggtatagcgaagctcaaacatgctgcatgcaacgcttgagtt 792
Db 664 GCGGATGCCGTGCGAGGCGATAGCGCAAGGCTCAAAACATTCGTCATGACGCGTTGGGT 723
QY 793 ctgcttccacccgaacaaagaatgacgcaacacacacacacacacacacacacacacac 852
Db 724 CTGCTTCCACCGCAAAACAGATGGCGCGCATCAACGATTTGGCGAGATTTGGCGCACTC 783
QY 853 aaagactatgcccgaagcagcaltccgcatgtggcagctccaaaccccaatgcccacaa 912
Db 784 AAAGACTATGCGCGAGCGACCATCGCGGATTTGGGAGTCCAAACCCCAATGCGCGCAAA 843
QY 913 ggcatagaagccgctcagcaatattcttattgacgacatcccatcaaaaggatggagct 972

Db 844 GGCATGAAAGCCCTGACATATCTTATGACAGCCATCCCATCAAGGATTGAGCT 903
QY 973 gtccgggaaataacggcttggcgagcatcaaggacatcctgtcaagcggttcgaatg 1032
Db 904 GTTCGGGAAAAATACGCTTGGGGGCGCATCAGCGACATCTTATCAACCGCTCGAGATG 963
QY 1033 ggcggatcgcatctggcgaaggaaatccggcgctcagcgacaatttgcgcatcgca 1092
Db 964 GGGCGATCCCATTTGGCGAAGGAAATCCGCCCTCAGCGCAATTTTCCGATCGGCA 1023
QY 1093 tacgcaaaataccgctccctaccatccctccgaataatccgttcaaatctggaagcgt 1152
Db 1024 TACGCCAAATACCGCTCCCTTACCATTTCCGAAATATCCGTCAACTTGAGAGCGCT 1083
QY 1153 tacggcaaaataacatcaacctcctcaacgctgcgcgcgttaaacggcaaaatgtcaa 1212
Db 1084 TACGGCAAAACAAATACACTCTCTCAACCGTCGCGCTCAAAAGGCAAAATGTCAAA 1143
QY 1213 ctggcaaacacgaccccccgaagcagcgctacggttgcggttaagggtttccgaat 1272
Db 1144 CTGGCAGACCAACGCCACCCGAAAGCAGCGGTACCGTTGACGTTAAAGGCTTCCGAAT 1203
QY 1273 ttggaagacgctgtaaatatgatacg 1299
Db 1204 TTTCGAAAGCAGTGAATATGATACG 1230

RESULT 16
AX236459 2256 bp DNA linear PAT 26-SEP-2001
LOCUS AX236459
DEFINITION Sequence 152 from Patent WO0164922.
ACCESSION AX236459
VERSION AX236459.1 GI:15796051
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 2256)
AUTHORS Arico,M.B., Comanducci,M.C., Galeotti,C.C., Maignani,V.C.,
Gulliani,M.M. and Pizzi,M.C.
TITLE Heterologous expression of neisserial proteins
JOURNAL Patent: WO 0164922-A 152 07-SEP-2001;
Chiron Spa (IT)
FEATURES
Source location/Qualifiers
1..2256
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="961c-ORF46.1"
BASE COUNT 698 a 608 c 545 g 405 t
ORIGIN

Query Match 63.3%; Score 1194; DB 6; Length 2256;
Best Local Similarity 98.3%; Pred. No. 1.9e-272;
Matches 1206; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 73 tcaagattggcaaaagatcccttatctcgcgaggttctgcacgctagatattcgaaacc 132
Db 1003 TCAGATTTGGCAAAAGATTCTTTATCCGGCAGGTTCTGACCGTCAAGCATTTGCAACC 1062
QY 133 gacggaaataccaccatctcgcgacgaggggagctctgcnaagcgcaacgcatatc 192
Db 1063 GACGGGAAATACACCTTATTCGGCAGCAGGGGGAACCTTCCGAGGCGACGCCCATATC 1122
QY 193 ggaattggaaacatacaagcatcagttggccaccgtatgaattcaacagcgccgctt 252
Db 1123 GGATTGGGAAAAATACAAAGCATCATGTTGGGCAACCTGATGATTCACAGCGCGCAT 1182
QY 253 gaaagaaatacgtgatactgtcgttttcgatacagcggaacaaattcgcacc 312
Db 1183 AAAGGAAATATCGCTACATTTGCTTCGCTTTCCGATCAAGGAGCAAGTCCATTCGCC 1242
QY 313 ttcgacaacatcgtctcaatctcgattctgacgaagccgtagtccggttgacgattc 372

Db 1243 TTTCGAAACCATCTCCACATTCCTGATTTCTGATGAACCCGGTATCTCCGTGACGATTT 1302
QY 373 agccttaccgcatcatttgggacggaatacgaacacatcccgccgacggtatgacgg 432
Db 1303 AGCTTTACCGCATTCATTGGGAGCGATGAGAACCATCTCCGCGAGGCTTATGACGG 1362
QY 433 ccacagggcgcgctatcccgctcccaaggcgcgaggaatalatacagtaacgata 492
Db 1363 CCACAGGCGCGCGCTATCCCGCTCCCAAGCGCGCAGGATATATACGTTACGACATA 1422
QY 493 aaagcgcttggcccaaaataatccgctcaaccttgacgcaacacgcaacgcaacgg 552
Db 1423 AAAGCGTTGCCCAAAATATTCGCGCTCAACCTGACCCACAAACCCAGACCGCAACGG 1482
QY 553 ctggcgaccggtttcccaaatggcgcgctatgctgtaacgaagagtaaggagcagatc 612
Db 1483 CTTCGCGACCGTTTCCCAATATGCGGTATGATGCTGAGCAGAGAGATGAGGACCGATT 1542
QY 613 aaacgcgcaacccgatalacagcccgagctggaacagatcgggcaatgcccgaagcctc 672
Db 1543 AAACGCCGCCACCGGATACAGCGCCGAGCTGAGATCGGCAATGCGCGCAAGCCTTC 1602
QY 673 aacggaactgcaagatcgtcaaaaacatcaatcgcgcgcaaggaataatgtcgcgca 732
Db 1603 AACGGCACTGCAGATATCTGTTAAACATCATCGCGCGGAGAGAAATGTGCGCA 1662
QY 733 ggcgattgcttgcaggttatgaagcgaaggtcaaacattgtgtcatgcaagcttgggt 792
Db 1663 GGGGATCCCGTGGAGGCGCATGACGCAAGCGTCAAAACATGCTGTATGATGACCGCT 1722
QY 793 ctgcttccacccgaaaacaaagatlgcgcgcatcaacgaattggcagataltggcgaactc 852
Db 1723 CTGCTTTCCACCGAAACCAAGATGCGCGCATCAACGATTTGGCGATATGCGCGCACTC 1782
QY 853 aaagactatgcggaagcagcagccatccggtatggcgagtcgaacaaacccatgcgcaaa 912
Db 1783 AAAGACTATGCGCGCAGCAGCATCCGCGATTGGGCGAGTCCAAACCCCAATGCCGACAA 1842
QY 913 ggcataagagcgcgcagaaatatttlaagcagcatcccatcaaaaggaattggagct 972
Db 1843 GGCATGAAAGCCGTCAGCAATATCTTTATGCGAGCATCCCATCAAAAGGATTTGAGCT 1902
QY 973 gtccgggaaataacggttggcgagcatcaagcagacatcctgtcaagcggttcgaatg 1032
Db 1903 GTTCGGGAAAAATACGCTTGGGGGCGCATCAAGCGCATCTTATCAAGGCGTGGCAGATG 1962
QY 1033 ggcgcatcgcatctgcgaaagggaatccgcgctcagcgacaatttggcgatgcgca 1092
Db 1963 GCGCGATCCCATTTGCCAAAGGGAATCCGCCGTACGCGACAAATTTGCCGATGCGCA 2022
QY 1093 tacgcaaaataccgctccctaccatccctccgaataatccgcttcaaatcttgagcagct 1152
Db 2023 TACGCCAAATACCGCTCCCTTACCATTTCCGAAATATCCGTTCAAACTTGGACAGCGCT 2082
QY 1153 tacgcaaaataacatcaacctcctcaacggttgcgcgcttcaaacgcaaaatgtcaaa 1212
Db 2083 TACGGCAAAAGAAACATACCTCTCAACCGTCCGCGCTCAAAAGCGCAAAATGTCAAA 2142
QY 1213 ctggcaagcaaacgcaacccggaagcagcggttacggttgaaggttaaggtttccgaat 1272
Db 2143 CTGGCAGACCAACGCGCACCCGAAAGACAGCGGTACGTTGACGGTAAAGGTTTCCGAAT 2202
QY 1273 ttggaagacgctgtaaatatgatacg 1299
Db 2203 TTTCGAAAGCAGTGAATATGATACG 2229

RESULT 17
AX239737 2256 bp DNA linear PAT 26-SEP-2001
LOCUS AX239737
DEFINITION Sequence 39 from Patent WO0164920.
ACCESSION AX239737

VERSION	AX239737.1	GI:15797379
KEYWORDS		
SOURCE	synthetic construct.	
ORGANISM	synthetic construct	
REFERENCE	artificial sequence.	
AUTHORS	1 (bases 1 to 2256)	
TITLE	Arico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M.	
JOURNAL	And Pizza,M.	
FEATURES	Hybrid expression of neisserial proteins	
SOURCE	Patent: WO 0164920-A 39 07-SEP-2001;	
	Chiron Spa (IT)	
	location/Qualifiers	
	1..2256	
	/organism="synthetic construct"	
	/db_xref="taxon:32630"	
	/note="ORF46.1-961c"	
BASE COUNT	698 a	608 c
ORIGIN	545 g	405 t

Query Match	63.3%	Score 1194;	DB 6;	Length 2256;
Best Local Similarity	98.3%	Pred. No. 1.9e-272;		
Matches 1206; Conservative	0;	Mismatches 21;	Indels 0;	Gaps 0;

QY	73	tcgagatttggcaaacggttcccttatcccggaagttcttgacgcctgaagtttcgaacc	122
Dp	4	TCGATTTGGCAAAACGATTTCTTTATTCGGCGAGGTTTCGACCGCTCAGCATTTTCGAACCC	63
QY	133	gaagcggaataataccaactctatccgycagcagggggaggtctgcacnagcgcacagccatatic	1920
Dp	64	GACGCGAAATATCCACTTATTCGGCGACGAGGGGGGAACCTTCCGACAGCGCACGGCCATATC	123
QY	193	ggaattggaaacatacaaaagccatcgaatttggcgaaccttgaatttcaaaagggggcgctt	252
Dp	124	GGATTGGGAAAATATCAAAACCCATTCATTGGGCAACTGATGATTCAACAGCGGCCATT	183
QY	253	gaaggaaataatcggctacatttgcgcttcttcgcatcagggcacaanaattccattcgcc	312
Dp	184	AAAGGAAATATCGGCTACATTTGTCGGCTTTCCGATCACCGGACAGCAAGTCATTCCTCCC	243
QY	313	ttcgacaacatgccttcaacattccgattcttgcagaagcggtatgtccggtttagacgattc	372
Dp	244	TTGAGACAACCATGCTCTACATTCCTCGATTCTGATGAAACCGGTATGCTCCGTTGACGGATTT	303
QY	373	agccttataccgcatccatttggagcaggtatagcaaacacatcccgccgcagcagctatgaacgg	432
Dp	304	AGCCTTTACCGCATTCATTGGGAGCGGATATGAAACACCATCTCCGCGACGGCTATTGACGGG	363
QY	433	ccacagaaggcgcgctatcccgcttcccaaaaggcgagaggtatatacagcttcgacata	492
Dp	364	CCACAGGGCGCGGCTATTCCTCCGTTCCCAAAAGCGCGAGGGATATTATACAGCTTAGCACTA	423
QY	493	aaagcgcttggccaaaatatcgcgctctaaccttaccgcagacaacccgcagcacgcgagacaacg	552
Dp	424	AAAGGGGCTTGGCCAAAATATTCGCTCTCAACCTGACCGACACCGCAGCAACCGGACAAACGG	483
QY	553	cttgcgacacggtttccacaaatgcccgcgtatgtctgaacgaagagtaagcgacgattc	612
Dp	484	CTTGCCGACCGTTCCTCACATATGCGGCTAGTATGTCTGATCCGAAGGATATAGCGCAGGATTTC	543
QY	613	aaacgcgcacaccgcatatacagcccgagcttggacagaatcggaatgcgcgcgaagccttc	672
Dp	544	AAACGGCGCACCCGATACAGCCCCGAGCTGGACAGATCGGGCATTCGCCCGAAGCCTTC	603
QY	673	aaacggaactgcgagatattgttcaaaaaacatcatccgcgcgcgcgcaggaagaaattgttcgcgca	732
Dp	604	AACGGCACTGCAATATTCCTTAAAAACATCATTCGGCGCGCAGAGAAATTTCTGGCGCA	663
QY	733	ggcagtcgcgtgcaaggtataagcgaaagctcaaacattgtcgtataagcagcgttgggt	792
Dp	664	GGGAGTGCCTGAGGGGCAATAAGCAGAAAGCTCAAACTATTCCTGTCTATGACAGGGCTTGGGT	723
QY	793	ctgccttccacgaaaaaagaatgtgcgcgcatcaaaagatttgcagataatgtgcgaacac	852

Db	724	CGGCTTTCACCCGAAACAAAGATGGGGCGCATCAAGATTGGGCAATATGGCGCAATC	783
OY	853	aaagcattagccgcagcaacccatccgcgatttggcagttccaaaaaccccaatgcgcgacaa	912
Db	784	AAAGACTTATGGCCGACAGCACCATTCGGCGATTGGGAGTCCAAAACCCCAATGCCGACAA	843
OY	913	ggcatagaagccgtaacgaatattctttagcagccatcccatcaaaaggatttgaagt	972
Db	844	GGCATAGAAGCCGTCAGCAATATCTTTATGGCAGCATCCCATCAAAAGATTGGAGCT	903
OY	973	gfcgggggaaaaatagcgtcttggcggaatcaagcaatccctgtcaagcggtccgaatg	1032
Db	904	GTTCCGGGAAATACGGCTTGGGCGGATCAACGGACATCTTATAAAGGGTGGGAGAG	963
OY	1033	ggcgcatctgcattgcggaagaagaatccgcgtcaacgaaatttggcgatgcgca	1092
Db	964	GGCGGATTCGCTATGCCGAAGAAGAAATCCGCCGTCAGGACAAATTTTGGCATGGCA	1023
OY	1093	taagcacaattaccggtccctcttaccattcccgaaatatccgttgaactttgaagaagct	1152
Db	1024	TACGCCAATACCCCTCCCTTACCATTCGCCAAATATCCGTTCAAACTTGGAGAGCGT	1083
OY	1153	taagcacaagaanaacatcaactcctcaacgcttgcgcggtcaaaacygcanaaatgttcaa	1212
Db	1084	TACGCCAAAGAAACATCACTCTCTCAACCGCGCGCTCAAAAGGCAAAAATGTCAAA	1143
OY	1213	ctggagagacaaagcacaacccgaaagacaagcgttaacggttttgcggtgaagaaggtttccgat	1272
Db	1144	CTGGCAGACCAAGCCGACCACCAAGCAAGCAGGCGTACCGTTTGACGCTAAAGGGTTTCCGAAT	1203
OY	1273	cttgaagaagcagctgaatatgatacg	1299
Db	1204	TTTGAGAGGACGTCGAATATGATACG	1230

RESULT	18		
AX239745			
LOCUS	AX239745	2256 bp	DNA
DEFINITION	Sequence 47 from Patent WO0164920.	linear	PAT 26-SEP-2001

ORGANISM	synthetic construct			
REFERENCE	artificial sequence.			
AUTHORS	1 (bases 1 to 2256) Alico,M.B., Comanducci,M., Galeotti,C., Masignani,V., Giuliani,M.M. and PizziA.M.			
TITLE	Hybrid expression of neiserial proteins			
JOURNAL	Patent: WO 0164920-A 47 07-SEP-2001; Chiron Spa (IT)			
FEATURES	location/Qualifiers			
source	1..2256			
	/organism="synthetic construct"			
	/db_xref="taxon:32630"			
	/note="g6ic-ORF46.1"			
BASE COUNT	698 a	608 c	545 g	405 t
ORIGIN				

Query Match	63.38	Score	1194	DB	6	length	2256	
Best Local Smilarity	98.38	Pred. No.	1.9e-272					
Matches 1206	Conservative	0	Mismatches	21	Indels	0	Gaps	0

Db	Qy	
1003	gaagggaataaccactatctgcgcagcagggggagcttcgcacgacgcaacggtccatc	192
1063	gacgggaatattccaccttattcggcagcaggggggaatttcgcagggcacggccaatatc	1122

[illegible]

QY	1620	taacaatgaggt-ttatcaagcgacatgtggaattaa	1655
Db	345	TAACAATGGGCTMTTATCAAGCGACAGTGAATAATTAA	381
RESULT	25		
AE002393/c			
LOCUS		11700 bp	DNA
DEFINITION			linear BCT 25-MAY-2000
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
M			

Db 7331 TCATTACGATTAAGCCCAACGAGCAGCACTAGAAACCGGCGGCAAAATACCACTCTTCG 7272
Qy 155 gcagcaaggaggagc---ttgcacnagcgcaagccatataatggttgaaacatacaaa 211
Db 7271 GCGACCGCGGCGAGCGCTTCCGACCGCACCGCAAAATCAACGATCCTCAAGACTATA 7212
Qy 212 gccatcagttgagccactgatatgattcaacaagcgcggttgaagaaatatacgctaca 271
Db 7211 CCCACGAGATGGGCAACTGCTCATCCAAAGCAAAATCAACGGCACATACGCGCTAC 7152
Qy 272 ttgtcgcgttttcgatacgaaggaacaaattccattcgccttcgaaacacatgctcac 331
Db 7151 ACACCCGCTTTCGAGACACGACAGCAAGAAACACGCCCTTCGACCAACGACGCGCG 7092
Qy 332 attcgcattcgaagaaacgagtagtcccgltgagcagattcagccttaccatcatt 391
Db 7091 ACAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7032
Qy 392 ggagcgaagtaacgaac 451
Db 7031 GGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6972
Qy 452 ccgcctcccaagcgagcgagatataatacagctacgacataaaggcggttgcacaaata 511
Db 6971 CCAAACTACGCGCGCGCGAGAGCAATACCTATCAGTCAACGCGCACAGCGCGAGTA 6912
Qy 512 tccgcctcaactacgacgaac 571
Db 6911 TCAAACTCAATCCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 6852
Qy 572 atgcgcgcctatgctgacgcaagagtagagcagcagattcaaacgcgcgacacacacata 631
Db 6851 ACCTGCGGAGCAATTTCTCCGACGCGCGGAGTAAGCCAAAGAAAATGTTGAGAGACA 6792
Qy 632 gcccgagcgttgacagatcgcggaatgcccgaagccctcaacgacgtcagatcg 691
Db 6791 ATGCCAATCGACGCGCGCGGAGCAACAGCATGATTTATCAACGCGCGCGCGCGCG 6732
Qy 692 tcaaaacatcatcgcgcgaggaagaattgctgcgcaagcagatgctgcgtgcaagta 751
Db 6731 CGCTCAACCCCTTTATTCAGCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6672
Qy 752 taaggaagcgtcaacaattgctgctatgacagcagcttgcgttccacacgaacaa 811
Db 6671 GCGCGTATGCGCATACAAAGCGCAATGCGCAATCGCCCTTGCCGCGAGAGGCA 6612
Qy 812 agatgagcgcaatcaagatttggcagatataggcgcaactcaagaactatgcccagag 871
Db 6611 AATTCGCGCTCATCGCGGCTTGGGCGAGCGCTGCGGCTTTGAAAAGAAATACGCGCAAG 6552
Qy 872 ccacgcgagattggcgagtcacaaaccccaatgcgacacagcagatagaagcgctcagca 931
Db 6551 CGGTGACCGGTGATACAGGAAATCCCAATGCGCGCGAAGCGCTGGAAGCGCTTCA 6492
Qy 932 atattctatggcgca 949
Db 6491 ACGTTCGCGCAGCAGCA 6474

RESULT 26
AX044029/c 349980 bp DNA linear PAT 24-NOV-2000
LOCUS AX044029 Sequence 108 from Patent WO0066791.
DEFINITION AX044029
ACCESSION AX044029
VERSION AX044029.1 GI:11342913
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis.
Neisseria meningitidis.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE
1 (bases 1 to 349980)
Plaza,M., Hickey,E., Peterson,J., Tettein,H., Venter,J.C.,
Maignani,V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M.,

Scarlato,V., Rappuoli,R., Frazer,C.M. and Grandi,G.
TITLE
Neisseria genomic sequences and methods of their use
JOURNAL
Patent: WO 0066791-A 108 09-NOV-2000;
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
Location/Qualifiers
1. 349980
/organism="Neisseria meningitidis"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces.-seq 1: 1 to
349980 349980 bases--seq 108: 300001 to 649980 349980
bases--seq 109: 600001 to 949980 349980 bases--seq 110:
900001 to 1249980 349980 bases--seq 111: 120001 to 1549980
349980 bases--seq 112: 150001 to 1849980 349980 bases--seq
113: 180001 to 2149980 349980 bases--seq 114: 210001 to
2272325 172325 bases"
BASE COUNT 84410 a 84863 c 94187 g 86520 t
ORIGIN

Query Match 15.0%; Score 283.4; DB 6; Length 349980;
Best Local Similarity 57.5%; Pred. No. 3.8e-56;
Matches 528; Conservative 0; Mismatches 387; Indels 3; Gaps 1;

Qy 35 ccatactggcagtgctgcctgcgcatgcatgcaacgcctcagattgcaaacgcatcct 94
Db 77278 CCGTAGGCGCGCGCGCACTATACAGCCCGCTCGCGCGGAGCTTGGCGCAAGACCGT 77219
Qy 95 ttatcgcgagttctcagcgtcagcatttcgaacacacgaaggaataacacattcgt 154
Db 77218 TCATTACCGATTAAGCCCAACGAGCAGCACTAGAAACCGGCGGCAAAATCAACGATCCTTTG 77159
Qy 155 gcaagcagggaggagc---ttgcacnagcgcaagcgcatactcagatttggaaacatacaaa 211
Db 77158 GGAACCGCGGCGGAGCGGCTTCCGACCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 77099
Qy 212 gccatcagttggcgaacctgatatgatacagcgcggttgaagaaatatacgctaca 271
Db 77098 CCCACGATATGGGCAACCTGCTCATCAACAGGCAAACTCAACGCGCAATCGGCTAC 77039
Qy 272 ttgtcgcgttttcgatacgaaggaagaattgctgcgcaagcagatgctgcgtgcaagta 331
Db 77038 ACACCCGCTTTTCGAGACGAGCAACGAGAAACAGCGCCCTTCGCAACGACGCGCG 76979
Qy 332 attcgcattcgaagaaacgagtagtcccgltgagcagattcagccttaccatcatt 391
Db 76978 ACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 76919
Qy 392 ggagcgaagtaacgaac 451
Db 76918 GGAAGGAGCAGCAACATCATCCCGCGATGCTACGACGCGCGGAGGCGGCAATTAC 76859
Qy 452 ccgcctcccaagcgcgagggatataatgattagcagataaagcgttgcacaaata 511
Db 76858 CCAAACTACGCGCGCGAGAGCAATACCTATCAGTCAACGCGCACAGCGCGAGTA 76799
Qy 512 tccgcctcaactacgacgaac 571
Db 76798 TCAAACTCAATCCGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 76739
Qy 572 atgcgcgctatgctgacgcaagagtagagcagcagattcaaacgcgcgacacacacata 631
Db 76738 ACCTGCGGAGCAATTTCTCCGACGCGCGGAGTAAGCCAAAGAAAATGTTGAGACACA 76679
Qy 632 gcccgagcgttgacagatcgcggaatgcccgaagccctcaacgacgtcagatcg 691
Db 76678 ATGCCAATCGACGCGCGGAGCAACAGCATGATTTATCAACGCGCGCGCGCG 76619
Qy 692 tcaaaacatcatcgcgcgaggaagaattgctgcgcaagcagatgctgcgtgcaagta 751
Db 76618 CGCTCAACCCCTTTATTCAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 76559
Qy 752 taaggaagcgtcaacaattgctgctatgacagcagcttgcgttccacacgaacaa 811

Db	76558	CGCGCTATGCATAGACAAAGCCGCAATGCGCAACATCGCCCTTGCCCGCGAGG6CA	76499
QY	812	agatggcgcatcaacgatttgcagatatgcgcaactcaagactatgccgcagcag	871
Db	76498	AATTGCGCGGTATCGCGCGCGTGGCAGCGCTTTGAAAGAAATACGCGCGAG	76439
QY	872	ccatcgcatctggcgatcccaaaccccaatgcagcagcagccgtcagca	931
Db	76438	CCGTTGACCGGTGGATACAGGAAAATCCCAATGCCCGCGAACCCTGAGCCGCTTCA	76379
QY	932	atattcttatggcagcca	949
Db	76378	ACGTTGCCGACGACGCA	76361
RESULT	27		
LOCUS	NMA722491	195767 bp	DNA linear ? RCT 04-DEC-2000
DEFINITION	Neisseria meningitidis serogroup A strain 22491 complete genome;		
ACCESSION	AL162758	AL157959	
VERSION	ALI62758.2	GI:7380672	
KEYWORDS	Neisseria meningitidis 22491.		
SOURCE	Neisseria meningitidis 22491.		
ORGANISM	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.		
REFERENCE	1 (bases 1 to 195767)		
AUTHORS	Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C., Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T., Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N., Holtzclow,S., Jagels,K., Leather,S., Moule,S., Mungall,K., Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M., Skelton,J., Whitehead,S., Spratt,B.G. and Barrall,B.G.		
TITLE	Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491		
JOURNAL	Nature 404 (6777), 502-506 (2000)		
MEDLINE	20222556		
REFERENCE	2 (bases 1 to 195767)		
AUTHORS	Parkhill,J.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-MAR-2000) Submitted on behalf of the Neisseria sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SA E-mail: parkhill@sanger.ac.uk		
COMMENT	Notes: Details of N. meningitidis sequencing at the Sanger Centre are available on the World Wide Web. (URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).		
FEATURES	Location/Qualifiers		
source	1. 195767		
gene	/organism="Neisseria meningitidis 22491"		
CDS	/strain="22491"		
	/db_xref="taxon:122587"		
	/note="serogroup: A"		
	80..952		
	/gene="NMA2045"		
	80..952		
	/note="NMA2045, possible oxidoreductase, len: 290aa; similar to many eg. SW:P20692 (TYRA_BACSU) prephenate dehydrogenase from Bacillus subtilis (372 aa) fasta scores; E(): 2.8e-30, 39.4% identity in 284 aa overlap."		
	/codon_start=1		
	/transl_table=11		
	/product="putative oxidoreductase"		
	/protein_id="CAB85263.1"		
	/db_xref="GI:7380673"		
	/db_xref="SPTREMBL:Q9J712"		
	/translation="MPTLNHIALIGVLLIGGSFVLDLKRGQIVRTVTGIDTDNDMLER ALERGVIDRASVAIDADISIGADLVLATPAVATPAITATLPVPEHTWISDVSTK SSYIEAFRCILPDRIRCIAPHSAGSDRSQAQAQGLPFRHKLITPHGGEHSGI ALIENMHAVGAETPTMDAQRHDAVFAAVSHMPLTATAVYHOUTLDHPDGCEYLKFAA TGPRDPTRIASGHPAWADICLANKSDLLDVGDLGKOLDVLAIDLITADDRALYRYF		

misc_feature	EBAKTTDRWDGN"	complement(347..556)	
misc_feature	/note="Core DNA uptake sequence: gccgcctgaa"		
gene	/label=DUS		
CDS	960..969		
	/note="Core DNA uptake sequence: gccgcctgaa"		
	/label=DUS		
	/complement(1009..2211)		
	/gene="NMA2046"		
	complement(1009..2211)		
	/gene="NMA2046"		
	/note="NMA2046, hypothetical integral membrane protein, len: 400aa; similar to TR:P95556 (EMBL:773914) hypothetical protein from Pseudomonas stutzeri (336 aa) fasta scores; E(): 0, 35.0% identity in 391 aa overlap. contains possible membrane-spanning hydrophobic regions."		
	/codon_start=1		
	/transl_table=11		
	/product="hypothetical integral membrane protein"		
	/protein_id="CAB85264.1"		
	/db_xref="GI:7380674"		
	/db_xref="SPTREMBL:Q9J711"		
	/translation="MIPFCGYINDKQTHMKFKHPVMAAFRPFSLAALYALSY ILMGRTGTHELSEFYHAEHIMGCVAGLVATFLTAATGQPPRGVYVGLT IFMLAARIAAFIPMGASASGILGTFPMYGAACMLPVRSGNORMYVAFFLVIG GTHAFHVLHNGNIGLISGOSGVYMGFGILGTRILSPTSRILVPOPSK VVAQASLMLPMLTALMHAGVWPWSAARPAAGVITTYOYTRMKRPVLEKPMWIL FAGVLEIGLIVAGASVFRPAFLNGLVHLDIGGIGVLTGMMARLALGHTGMPILP PKAVPAVFMMAATAVRMVAVFSSGTATYHSIRTSVFLALALVYAKYIPMLIR PRSDRPc"		
	complement(2208..2576)		
	/gene="NMA2047"		
	complement(2208..2576)		
	/note="NMA2047, hypothetical protein, len: 122aa; weakly similar to SW:P71932 (Y003_MCTU) hypothetical protein from Mycobacterium tuberculosis (161 aa) fasta scores; E(): 0.017, 32.5% identity in 117 aa overlap."		
	/codon_start=1		
	/transl_table=11		
	/product="hypothetical protein"		
	/protein_id="CAB85265.1"		
	/db_xref="GI:7380675"		
	/db_xref="SPTREMBL:Q9J710"		
	/translation="MKPLKHPALIGLSRDHHSLSLCVRLTPTPERHNDLEPHPS ELTTHFRREEAKFAITMONVARELQRFKDHARLROMASPRYGNAAWTAFTTLR DHAREERELPVAEPFLPA"		
	complement(2622..2631)		
	/note="Core DNA uptake sequence: gccgcctgaa"		
	/label=DUS		
	2750..3193		
	/gene="NMA2048"		
	2750..3193		
	/note="NMA2048, conserved hypothetical protein, len: 147aa; similar to many eg. SW:P21498 (YJER_ECOLI) hypothetical protein from Escherichia coli (141 aa) fasta scores; E(): 5.9e-23, 47.9% identity in 140 aa overlap."		
	/codon_start=1		
	/transl_table=11		
	/product="conserved hypothetical protein"		
	/protein_id="CAB85266.1"		
	/db_xref="GI:7380676"		
	/db_xref="SPTREMBL:Q9J709"		
	/translation="MYTLQHDYDGRVLIYTAINDADVNSITAVYTGSKHMKV VTAIVKGFSLHSVKGKGLRIASPERRNITCAVVRHNEPMDVCEGMENNECLTPS CRTLIGIGAMKSFPTVLDGFTLDLKLKPYDLLEPRIPAVQ"		
	3275..3314		
	/note="stem loop containing DNA uptake sequences: gcaat gccctctgaa cagccgatgc ttcaagcgc atcgcg"		
	3280..3289		
	/note="Core DNA uptake sequence: gccgcctgaa"		
	/label=DUS		

ACCESSION AE002420 AE002098
VERSION AE002420.1 GI:7225876
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
1 (bases 1 to 10057)
Tetelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
Cittone, H., Clark, E.B., Cotton, M.D., Uitterback, T.R., Khouril, H.,
Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Maignani, V.,
Pizza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappelli, R. and Venter, J.C.
Complete genome sequence of *Neisseria meningitidis* serogroup B
strain MC58
Science 287 (5459), 1809-1815 (2000)
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
2 (bases 1 to 10057)
Tetelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
Cittone, H., Clark, E.B., Cotton, M.D., Uitterback, T.R., Khouril, H.,
Qin, H., Vamathevan, J., Gill, J., Scarlato, V., Maignani, V.,
Pizza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappelli, R. and Venter, J.C.
Direct Submission
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
LOCATION/Qualifiers
1. 10057
/organism="Neisseria meningitidis MC58"
/strain="MC58"
/db_xref="taxon:122586"
/note="serogroup: B"
126. .350
/gene="NMB0649"
126. .350
/gene="NMB0649"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAFA1070.1"
/db_xref="GI:7225877"
/translation="MRLATKDLAIAIRKGVKSSPNEQLRAIKGESKIPDYTMHH
HODTGRMOLIREGLHDTGHIIGWAMNKR"
323. .742
/gene="NMB0650"
323. .742
/gene="NMB0650"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAFA1071.1"
/db_xref="GI:7225878"
/translation="MGSNQRKVTMMKLIKEDSDLEPAIKCIFSQSIDLNEFKIMTE
VYIRKMPIDIPYIDLPDGDGIDINIVGVSSCRKSKKNALIGLAFKGLID
VVDPLSKAKALKAEKHEPITYQHFHFPELPL"
843. .1232
/gene="NMB0651"
843. .1232
/gene="NMB0651"
/note="hypothetical protein; identified by Glimmer2;
putative"

putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAFA1072.1"
/db_xref="GI:7225879"
/translation="MKLIKEDSDLDGFAIKCLFQSISIDINEFKLMIEQYIRDPED
IPFYIFDLADPDGIGADINIVGVSSCRKSKKNALIGLAFKGLIDVYDPLSKKEK
ALKAEKHEPITYQHFHFPELPL"
1355. .2296
/gene="NMB0652"
1355. .2296
/gene="NMB0652"
/note="similar to GP:4838554 percent identity: 79.35;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="mafa protein"
/protein_id="AAFA1073.1"
/db_xref="GI:7225880"
/translation="MKTLLILPIVLTACGTLTGIPAHGGRKRPAYEDELVAASRAA
VKEMDLISAIGKRAALYVSMGDSGSGISDALIRGCYHNHPESATQSYTA
YDTATRTSDALSSVTTSTSLNAPAAALFKSGKGENSAGLSVNGTGRNETLIA
NPQVSLFNLIQIVFLGIEVVPEDVTVFVADVFVGFVRSFTEHLIYAEETK
AOTRIEYFVAVDRDSKRLITPKTAAYESQYOEYALMTGPKYKSKVYKASDRIMVDFS
DITPYGDTAANRPDPKONNGKKPDGVNEVIRRKGG"
2300. .3568
/gene="NMB0653"
2300. .3568
/gene="NMB0653"
/note="MafB-related protein; identified by Oxford;
putative"
/codon_start=1
/transl_table=11
/product="MafB-related protein"
/protein_id="AAFG2314.1"
/db_xref="GI:7413436"
/translation="MKPLRLNLAAACVAAAALIQPALADADDPFTTDAOROH
YEPGKYHLFEDPRGSVSDRTGKINVITYHQMGNLIIQANINGTIYHTFSGHG
HEENAPFHNHAADSASEKGNVDEGTYVRLMGEHNPADADYDPKGNPKPTGA
RDEYTHYNGTARSTIKLNPDTDRSIIRORISDNYSNESPDADANRPMFHNKTL
DRKNSMFEINGVAAGALNPISAGEANDQNMQENNAETVALYVLPFAVKNIT
KAAPGKAASGDSSTYKHNATSRISOSVDGEMPTRVNDFRAKSIKIHGDAQOK
HISGRNITEGKSTLQNIINPELLNGHSGAVPYISKGARGNPVDFGPIGSDGKS
GLSTNFGTHSGKNGVHIVPANPKTIKKVQ"
3571. .3888
/gene="NMB0654"
3571. .3888
/gene="NMB0654"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAFA1074.1"
/db_xref="GI:7225881"
/translation="MNIPLSWLRVGMNIAIMGLIHSIDRLITVDEGRERFLIKNVL
SREAITDEHDEMELYLTLMWSWCGEYFPEADRECIYSNHSSMELNGLNGAVFRKKELI
SQA"
4023. .4730
/gene="NMB0655"
4023. .4730
/gene="NMB0655"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAFA1075.1"
/db_xref="GI:7225882"
/translation="MVKTADCYKAKARIPTGDRVYAKDEASGKTKPYTARYGNPYQ
ETVYIETSDIGNQTILSNKIHPIYSGKMWIQAQRLKKGDTLSSGAKQYVONITF
KOPLKAYNLTIVADWHYFVKGSQAETGEVWVHNDCPYDKGNORYKDAASYHGKNDNSV


```

misc_feature
/note="Core DNA uptake sequence: gccgctcgaa"
/label=DUS
complement(1116..2030)
/gene="nuoL"
/note="Pfam match to entry PF00361 oxidored_q1,
NADH-ubiquinone/plastoquinone (complex I), various chains,
score 351.80, E-value 7.3e-102"
complement(2061..2243)
/gene="nuoL"
/note="Pfam match to entry PF00662 oxidored_q1_N,
NADH-ubiquinone oxidoreductase (complex I), chain 5
N-terminus, score 77.20, E-value 3.5e-19"
complement(2487..2828)
/gene="NMA0003"
complement(2487..2828)
/note="NMA0003"
/codon_start=1
/transl_table=1
/product="hypothetical protein NMA0003"
/protein_id="CAB83323.1"
/db_xref="GI:7378781"
/db_xref="SPTREMBL:O9JU91"
/translation="MDKNLEDFHAKMLNITYEALNPPIRYATRLQIVNEFGGK
EADAKLSTGEKKTQGTFTELLISGGVGHALKYMEYLVLOKPCWCDLFTTEQLAVARK
RLRVGFVSPK"
complement(2833..2837)
complement(2952..3527)
/gene="NMA0004"
complement(2952..3527)
/note="NMA0004, unknown, len: 191 aa; similar to
hypothetical proteins e.g. Y977_HAEIN HI0977 (191 aa),
fasta scores: E(): 0; 80.5% identity in 190 aa overlap,
and to proteins involved in cell cycle e.g. TR:O54679
(EMBL:AF036487) Lactococcus lactis Plasmid pNZ4000
putative mobilization protein (200 aa), fasta scores: E():
0 55.1% identity in 167 aa overlap, and FIC_ECOLI cell
filamentation protein FIC (200 aa), fasta scores: E():
0.093, 29.4% identity in 126 aa overlap"
/codon_start=1
/transl_table=1
/product="hypothetical protein NMA0004"
/protein_id="CAB83324.1"
/db_xref="GI:7378782"
/db_xref="SPTREMBL:O9JU09"
/translation="MPSENPIGTMTKSIDEOSLHNAARLESGDIDRIEVTAGLQO
IHRYLFGGLYDFAGQIRENISKSGFRFAAMLYKELVLEIOMKEPFEELIKYYE
KMDIAHFLLENGSTRITWLDLVKKNLKKVNMVNONSKITLYLQAMERSPVNDLEKRL
LNDNLTDVDNREILFKIGTIOGYEYENG"
complement(3516..3525)
/gene="NMA0004"
/note="Core DNA uptake sequence: gccgctcgaa"
/label=DUS
complement(3561..3866)
/gene="nuoK"
complement(3561..3866)
/gene="nuoK"
/EC_number="1.6.5.3"
/note="NMA0005, nuoK, NADH dehydrogenase I chain K, len:
101 aa; similar to many e.g. NUOK_RHOCA NADH dehydrogenase
I chain K (EC 1.6.5.3) (102 aa) fasta scores: E():
1.1e-24; 65.0% identity in 100 aa overlap. Contains Pfam
match to entry PF00420 oxidored_q2,
NADH-ubiquinone/plastoquinone oxidoreductase chain 4L"
/codon_start=1
/transl_table=1
/product="NADH dehydrogenase I chain K"

```

```

misc_feature
/protein_id="CAB83325.1"
/db_xref="GI:7378783"
/db_xref="SPTREMBL:O9JU09"
/translation="MTLTHYLVGLALFLGISANGIPMNRKNVLLMSIELMLAVN
FNFIASFQHLGDTYAGQIFVFVFLVVAESAIGLAIWLVYRNQITINVDLEKQ"
complement(3564..3863)
/gene="nuoK"
/note="Pfam match to entry PF00420 oxidored_q2,
NADH-ubiquinone/plastoquinone oxidoreductase chain 4L,
score 141.50, E-value 1.5e-38"
complement(3863..4534)
/gene="nuoJ"
complement(3863..4534)
/gene="nuoJ"
/EC_number="1.6.5.3"
/note="NMA0006, nuoJ, NADH dehydrogenase I chain J, len:
223 aa; similar to many e.g. NUOJ_RHOCA NADH dehydrogenase
I chain J (EC 1.6.5.3) (202 aa), fasta scores: E():
9.7e-21; 41.3% identity in 201 aa overlap. Contains Pfam
match to entry PF00499 oxidored_q3,
NADH-ubiquinone/plastoquinone oxidoreductase chain 6"
/codon_start=1
/transl_table=1
/product="NADH dehydrogenase I chain J"
/protein_id="CAB83326.1"
/db_xref="GI:7378784"
/db_xref="SPTREMBL:O9JU90"
/translation="MTFOLLIFYIFVNIILYGAIKYTAKNPVHAHLVLTFCVSAM
LMLMOAEFLGYLVVVYGVAVVLEFVVMNIDIEERAGFWRHAPVAGVGTLL
AVALLILVNPRTDLAAGFIMKDIADYNNI ROLGSIYIDYLLPFLAVLLLGAV
AAIYALVRRKTVNPKRMDPADQYVVRADQGMRLVYKMEAVAPQESAESEVSDDLKRC
ERGA"
complement(3874..3879)
/gene="nuoJ"
complement(4034..4525)
/gene="nuoJ"
/note="Pfam match to entry PF00499 oxidored_q3,
NADH-ubiquinone/plastoquinone oxidoreductase chain 6,
score 77.90, E-value 2e-19"
complement(4570..5277)
/gene="NMA0007"
complement(4570..5277)
/gene="NMA0007"
/note="NMA0007, unknown, len: 235 aa"
/codon_start=1
/transl_table=1
/product="hypothetical protein NMA0007"
/protein_id="CAB83327.1"
/db_xref="GI:7378785"
/db_xref="SPTREMBL:O9JU89"

```

Query Match 5.7%; Score 108.2; DR 1; Length 340806;
Best Local Similarity 62.3%; Pred. No. 1,4e-14;
Matches 170; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

```

OY 239 aacagcgagcgcttgaagaatatacgtctacattcgcgcttccgatacagcgagaca 298
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 305685 AACGGACAGATTGGAAGCATATTGATTAAGAACCCATTTTCAGGACATGACATG 305626
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 299 aattccatcgcccttcgaacaacatgcctacattcgcgatttcgagaaagcgtagtc 358
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 305625 AAGTACACAGCGTCGTTGATATCATGATTCAMAAAGCACTTCTGATTCAGCGGCGCG 305566
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 359 ccgttgacgagcttcaagccttcgcgatcatattggagagatacgaaccatccgcgcg 418
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 305565 TAGACGGTGCGTTTACCGTTTACCACTTCATCGGACAGGGTCGGAATTCACGAGG 305506
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 419 acggtcatgacgagcacacagcgagcggtatccgcctcccaagagcgagcgagatata 478
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 305505 ATGGAATATGACGCGCGCAAGGACGATTATCCGCCCCCGGAGAGACAGGATATAT 305446
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 479 acagctacgacataaaggcgcttgcaccaata 511
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 305445 ACAGCTACTATGTCACGAAGCACTTCACACAAAA 305413

RESULT 34
A96134 1449 bp DNA linear PAT 07-SEP-2000
LOCUS Sequence 167 from Patent WO9924578.
DEFINITION A96134
ACCESSION A96134.1 GI:6779959
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);
CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)
FEATURES
SOURCE
1. 1449
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 445 a 310 c 335 g 340 t 19 others
ORIGIN

Query Match 5.6%; Score 105.2; DB 6; Length 1449;

Best Local Similarity 61.2%; Pred. No. 3.3e-14; Matches 167; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 239 aacaggcgccgcttgaaggaatatcgctacatctgctccgtttccgcatacagggcaca 298

Db 245 AACGGACAGGATTGAAAGGTTATCGGTTATGAACCATTTTTCAGACATGACGACATG 304

QY 299 aattccattcgcccttcagacacacatcctccatcattcgatttcgaagccggtagtc 358

Db 305 AAGTACACAGCCGTTGATCATGATTCACAAAAGCACTTCTGATTCAGCGCGCGC 364

QY 359 ccgttgaaggtatcgcttccgcattccatcattggaagcagatacgaaccatcccgccg 418

Db 365 TAGACGGGTGTTTACCGTTTACCAACTTCATCGACAGGGTCGAAATCCATCCGAGG 424

QY 419 acggtatgaagcgccacagggcgcggtcgtccctcccaagggcgcgaggtatata 478

Db 425 ATGGATATGACGGCGCCGACAGCAGCATATTCGCCGCCCGGAGAGAGACAGGATATAT 484

QY 479 acagctacacataaaggcggtgccccaaata 511

Db 485 ACANTTANTATGTCAAGCACTTCACACAAAA 517

RESULT 35
A96132 1452 bp DNA linear PAT 07-SEP-2000
LOCUS Sequence 165 from Patent WO9924578.
DEFINITION A96132
ACCESSION A96132.1 GI:6779958
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);
CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)

FEATURES
SOURCE
1. 1452
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 466 a 298 c 332 g 356 t

ORIGIN

Query Match 5.6%; Score 105; DB 6; Length 1452;

Best Local Similarity 61.3%; Pred. No. 3.6e-14; Matches 168; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 239 aacaggcgccgcttgaaggaatatcgctacatctgctccgtttccgcatacagggcaca 298

Db 245 AACGGACAGGATTGAAAGGTTATCGGTTATGAACCATTTTTCAGACGACGACATG 304

QY 299 aattccattcgcccttcagacacacatcctccatcattcgatttcgaagccggtagtc 358

Db 305 AAGTACACAGCCGTTGATCATGATTCACAAAAGCACTTCTGATTCAGCGCGCGTG 364

QY 359 ccgttgaaggtatcgcttccgcattccatcattggaagcagatacgaaccatcccgccg 418

Db 365 TAGACGGCGGTTTACCGTTTACCAACTTCATCGACAGGGTCGAAATCCATCCGAGG 424

QY 419 acggtatgaagcgccacagggcgcggtcgtccctcccaagggcgcgaggtatata 478

Db 425 ATGGATATGACGGCGCCGACAGCAGCATATTCGCCGCCCGGAGAGAGACAGGATATAT 484

QY 479 acagctacacataaaggcggtgccccaaata 511

Db 485 ACAGCTATATGTCACGAAGCACTTCACACAAAA 517

RESULT 36
AE002559 10869 bp DNA linear BCT 25-MAY-2000
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

1 (bases 1 to 10869)
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., Deboy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
Cittone, H., Clark, E.B., Cotton, M.D., Uterback, T.R., Khouli, H.,
Olin, H., Yamahewyan, J., Gill, J., Scarlato, V., Masignani, V.,
Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappuoli, R. and Venter, J.C.
Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58

Science 287 (5459), 1809-1815 (2000)

10710307

2 (bases 1 to 10869)

Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., Deboy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
Cittone, H., Clark, E.B., Cotton, M.D., Uterback, T.R., Khouli, H.,
Olin, H., Yamahewyan, J., Gill, J., Scarlato, V., Masignani, V.,
Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappuoli, R. and Venter, J.C.
Direct Submission

Submitted (17-Mar-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
On Apr 4, 2000 this sequence version replaced gi:7227357.

Location/Qualifiers

1. 10869
/organism="Neisseria meningitidis MC58"
/strain="MC58"

BASE COUNT 466 a 298 c 332 g 356 t

```
gene
CDS
/db_xref="taxon:122586"
/note="serogroup: B"
150..524
/gene="NMB2095"
150..524
/gene="NMB2095"
/note="similar to PID:1817528 percent identity: 53.99;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="adhesin complex protein, putative"
/protein_id="AA042412.1"
/db_xref="GI:7227358"
/translation="MKLTFAILSSAIALSSMAAAGTDNPTVAKKTVSYCGCKKV
KVIYGENKOGITVYASAVINGKRVOMPYNLDKSDNVEFTYEGEGYVIGTGYMDKSY
RKOPITAPDNOIVFKDCSPR"
complement(712..2178)
/gene="NMB2096"
complement(712..2178)
/note="similar to SP:P33940 PID:424051 GB:U00096 percent
identity: 69.73; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="malate:quinone oxidoreductase"
/protein_id="AA042413.1"
/db_xref="GI:7227359"
/translation="MAEATDVILVGGIMSAITGILLKELEPSMEITLIERLEDALE
SSNANNAGTGHSCALCELNAPLGANGIIDPARLNTAEQHVSRQFWATLVAEGKLE
RDNQVPAANTSAEGTVDGRLTRQWVKYLGQGVTEFNHVEDIKRESDGAVLTK
TADTNPDDQLTTRFLFLGAGGALTLQSGIIPGKGYGFPVSGLFPNSNPET
AECHNAKYGOASGAPPMGVPPIHLDTRNVGKRHLMEGPGVAGFSPNFKOGSLMDPL
SIHMDLYPMLCAGMANMPLTKYLGLFKTKKEERPAISLEYPEANDDMELITTAGO
RVQIITKDSKGVIGORGETIYVAHAGDSIALGASGASIAVLMIRLMQCPEPRA
PSWEDRLKEIYPGYGIKLMENPERADITIAVTAYLDI"
complement(2230..2319)
/gene="NMB2097"
complement(2230..2319)
/gene="NMB2097"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA042414.1"
/db_xref="GI:7227360"
/translation="MLATHSNPPTAVSDTHHCRFSLLRRRL"
complement(2715..3305)
/gene="NMB2098"
complement(2715..3305)
/gene="NMB2098"
/note="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA042415.1"
/db_xref="GI:7227361"
/translation="MSEKILIPVLSIGVRLPELDVHHTGLREAVDGEVWKIGYVS
APHPRVADYIGTATATRLAFVDEEAGRYVTTAYTHPEQIPIRLDIGFTWASAA
RRTIRNTCKKMLLDSAFVLCRCVGRWDILNLASQRIERLGAERDGLRHMHLR
KQGSYRDIVVYSMLREDCKNREILGRLAGYGVY"
complement(3366..3947)
/gene="NMB2099"
complement(3366..3947)
/gene="NMB2099"
/note="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/gene
CDS
/protein_id="AA042416.1"
/db_xref="GI:7227362"
/translation="MRNEERKALRRELGRRSQMGROVRAATYKINHLIRYIKGR
KIGVYWMGKELRLDGFVRAAQKRGALYIPYIPRSRMFTYYPAGVGOEKRGR
AKLHVPOFAGRKRRKRDNLILVYVGMDRGLGYLGAGGYDATTLSMKKRLQAKTV
GVGFACQIVDRLPVEAHDRSLDGFVSNGLICF"
4025..4258
/gene="NMB2100"
4025..4258
/gene="NMB2100"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA042417.1"
/db_xref="GI:7227363"
/translation="MMHIVAIGYLFVAVMYSAQPSIARALILYFVNAVLPYFTVFT
ITVRRNHLMRQEQAESBQQRARQKDSGTFK"
4431..5159
/gene="NMB2101"
4431..5159
/gene="NMB2101"
/note="similar to SP:P34831 GB:X53651 PID:296030 percent
identity: 78.02; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="30S ribosomal protein S2"
/protein_id="AA042418.1"
/db_xref="GI:7227364"
/translation="MSQIMROMIEAGVHGHOTFRPMNPKMAOYIRGARKIHVNE
KILPMQDQAEVRRLVARKGVTLFVSTKROADITREETFRAGMPVDVIRWIGMIT
NKTYKQSRILEKTAALENAESESPSKETILEMDQVEKLSRISGILNMGLDPA
ITVIDTGKGLVVEAKELGPIYAVVDVNNSPDGKYYVIPGNDSSAKAIRLYCRGIA
DVLEKGNQALDETVAADQEAEE"
5288..6142
/gene="NMB2102"
5288..6142
/gene="NMB2102"
/note="similar to GB:D26562 SP:P02997 GB:D13334 GB:V00343
PID:1208943 percent identity: 69.01; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="elongation factor TS (EF-TS)"
/protein_id="AA042419.1"
/db_xref="GI:7227365"
/translation="MAETIAKVVADLRAATGLGMECKKALVEAGNFDKAEILIRIK
SGAKAGKLAGRTAAGVLAIVAINGVGLVEVCEIDFVAKDGFVEAFVAKTAAE
KKPASVEELSELVEAEKRAIIAKLGEMSVRRVOYIDFNQDLVAYIHGALATGVLVE
YKGSDEVARIKIGMHIYAAKPOCVSEVNAEYFEKRRIYTBQALIASGPAQDIAAKV
EGRIKPLAETILNQAPFMPNDQYVAPSKENGTEYISFVRKIKVDGIEKKAVDYAA
EVAARAKV"
6359..7078
/gene="NMB2103"
6359..7078
/gene="NMB2103"
/note="similar to GB:D26562 SP:P29464 GB:D13334 GB:X78809
PID:1208944 percent identity: 73.16; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="uridylylate kinase"
/protein_id="AA042420.1"
/db_xref="GI:7227366"
/translation="WTQQIKYKRVLLKSGESIMGSDPGINHDTIVQTVGEIAEYVK
MGVOVGIVGGGNIIRGVSAQGSMDRATADGMATVMNALALADAEFTGIGIKRV
OSALSMOQIAETVYAPKAIQYILEGKVYIFAGTGMPFTPTTAALRGAENQDML
KATNVDDGYTADPKRDPSATRYETITPEALILKNKLVMDATFALCREKKNIVYFGI
AKESIKRVITGEGDEGTIVHC"
7262..8225
/gene="NMB2104"
```


LOCUS	166494	7218 bp	DNA	linear	PAT 28-DEC-1997
DEFINITION	Sequence 14 from patent US 5670367.				
ACCESSION	166494				
VERSION	166494.1	GI:2724471			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 7218)				
AUTHORS	Dorner,F., Schefflinger,F. and Falkner,F.Gunter.				
TITLE	Recombinant fowlpox virus				
JOURNAL	Patent: US 5670367-A 14 23-SEP-1997;				
FEATURES	Location/Qualifiers				
SOURCE	1..7218				
BASE COUNT	1944 a	1491 c	1486 g	1929 t	368 others
ORIGIN					
Query Match	4.3%; Score 80.2; DB 6; Length 7218;				
Best local Similarity	5.5%; Pred.No.3.5e-08;				
Matches	22; Conservative	239; Mismatches	142; Indels	0; Gaps	0;
OY	1357 gatgcgaacccgagatggagggtgatagaacttaataattgacaactcgtgaacag				
Db	1424 RR				
OY	1417 gtggagaaaagtctcgagaaacgagaagagtgatcgagtgaatttaaagcccat				
Db	1364 RR				
OY	1477 gcgcacagcaaatgggaaaaataaacagagttagattaatcatctttaagtcgtgat				
Db	1304 RR				
OY	1537 atcataagaagaagcacagtaacagggagcatagcttcaaccgtggtatgtacgggtg				
Db	1244 RR				
OY	1597 atacaacaacctgcgcacctgataaactgttgttacagagcagatggaaattaa				
Db	1184 RR				
OY	1657 aagcccgatggaagtgtgggggtgaaaaacgaaaaaggtgggaagtgtgaccaagcac				
Db	1124 RR				
OY	1717 accatgtccccaaagatgtggatgcagctgaatttagcgctg	1759			
Db	1064 CGCAAGCTCCTCGACTGCAGCCAACTCGGATTATTTGTG	1022			
RESULT 40					
AF429315					
LOCUS	AF429315	125020 bp	DNA	linear	PRI 18-JAN-2002
DEFINITION	Homo sapiens junctophilin 3 (JPH3) gene, partial cds.				
ACCESSION	AF429315				
VERSION	AF429315.1	GI:17646244			
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 125020) Holmes,S.E., O'Heern,E., Rosenblatt,A., Callahan,C., Hwang,H.S., Ingerson,I.Ashworth,R.G., Fleisher,A., Stevanini,G., Brice,A., Potter,N.T., Ross,C.A. and Marzolis,R.L. A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2 Nat. Genet. 29 (4), 377-378 (2001)				
TITLE					
JOURNAL					
MEDLINE					
PUBMED	11694876				
REFERENCE	2 (bases 1 to 125020)				

FEATURES	source
<p>AUTHORS Holmes, S.E., Ingersoll-Ashworth, R.G., Ross, C.A. and Mergolis, R.L.</p> <p>TITLE Direct Submission</p> <p>JOURNAL Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA</p> <p>Location/Qualifiers 1. 125020</p> <p>organism="Homo sapiens"</p> <p>db_xref="taxon:9606"</p> <p>chromosome="16"</p> <p>map="16q24.3; between D16S20 and WI-12410"</p> <p>note="Isolated from a patient with Huntington's disease-like 2 (HDL2)"</p> <p>complement (35581..35746)</p> <p>fp1_type=tandem</p> <p>fp1_unit=ctg</p> <p>complement (<36507..>36887)</p> <p>gene="JPH3"</p> <p>product="junctophilin 3"</p> <p>complement (<36507..>36887)</p> <p>gene="JPH3"</p> <p>note="Jp3"</p> <p>complement (<36507..36887)</p> <p>gene="JPH3"</p> <p>note="component of the junctional complex between plasma membrane and endoplasmic reticulum"</p> <p>codon_start=1</p> <p>product="junctophilin 3"</p> <p>protein_id="AAL40941.1"</p> <p>db_xref="GI:17646245"</p> <p>translation="MSSGGRNPFDDGGSGCGMEDGKAHGHGVTGPRKGQGEYGSMSHGFEVLGVTPMSGNTYGVGIWAQGRKHGIGLESGKWWYKGEWTHGFKGRGVRECAAGNAKELEGVWSMGLDDGVTETYSNG"</p>	<p>BASE COUNT 29056 a 32731 c 30659 g 28283 t 4254 others</p> <p>ORIGIN</p>

Query Match	2.6%	Score 49	DB 9	Length 125020
Best Local Similarity	10.5%	Pred. No. 1.3		
Matches 96	Conservative 395	Mismatches 419	Indels 8	Gaps 3
QY 314	tcgacaacacatccccaacatcttcgatttcgacgaagccgagatgccggtatgccggtatgaacgattca	373		
Db 16975	KMRBSMNCSTSCYASMCMMCCMSGCCMRSCCMCRSTCCMRCCACRYCCACKCYSSYWTMSA	17034		
QY 374	gccttaccgcatcccatcttgagagagatagaacaacatcccgacgagcatatgagc 433			
Db 17035	SYMRSRYWRKSMKCMWSMCSRSRSKCSNKGCGSMKKGKGSYGRKTKRKSRGKMG	17094		
QY 434	cacagagcgagcgagctatccgctcccaaggcgcgagagatatacaagctacgacataa	493		
Db 17095	AKWMYTYRRSMKMKMYSSKGMVYCMCMWRRCGYCSGMTSRAMCCSYACRCKSMCY	17154		
QY 494	aagcgctgcccacaanlctgcctcaacccgagacgcgaacccgacagacgcgaaagcgc	553		
Db 17155	SCYYSMSKGYTYACSYGSMSSRYCMNGSYSTSCGCCCTTTTCCCCNANTGGGAA	17214		
QY 554	ttgcgcagcttccacatgcgcgcgcgtacgtcagcaagaagatlagcgacgattca	613		
Db 17215	GCTTTTNCNKRTYSYRKRNCAMCKYNNYNNNSBRSCRGMSCTRYKSSMTMSAStc	17274		
QY 614	aacgcgcaccccgatatacagcccgagctggaacagatcgcggcaatgctgcgcgaagcttca	673		
Db 17275	WCMNMYCYSMKSMASRGMSWSYMKMKKSSRMSCYMKCCMCKMCCYMRSMRSRMSY	17334		
QY 674	aaggcagctcagagatctgcataaacaatcctgcgcgcgcagagagaagaatgtgcgcgcag	733		
Db 17335	YMYSMKSSSSNGCCTRCYCMSSKSCYKSYMMMRKRMKMKWSGMSWGMSASRSS	17394		
QY 734	gcgactgcgctcagagatagaacgaaagctccaacatctgtctcaty--caagcttgagg	791		
Db 17395	CYKCYKSMRCSMSSKCYRCAGCMMMGVIMRYRYRCHSMKRRMKGSAMTBMRRMKGRG	17454		
QY 792	tcgtcttccacgcgaaaaaagaatgycgcgcgcataacgaattgtgcagatattgycgaact	851		

Db 17455 AMAMCKCYRMRSCMMAMSYCASCRSCAMSGTYMCKASYMCKSSYWCCTWCSTSYMSY 17514
Qy 852 caaagactatcgccgagcagccatccgcatgtgg---cagtcocaaaccccaatgcgcg 907
Db 17515 YRCTCMCKGMSYTKSKMSWSSMSYKKRGRKYSMTCTSGAMSCMRCCYMRGASSMR 17574
Qy 908 cacaagcagtagaagccgtcagcaatattttagcgagccatcccatcacaagagatg 967
Db 17575 AGSMRRRAAGRSWGGGSKMMWTGMMRSKYYYCTGRMMATTYMC--CWRRRSYMYR 17632
Qy 968 gaagctccgggggaaatacagcttgagcgccatcagcagcagcatcttcaagcgctgc 1027
Db 17633 SMSMGMRKSSWSGMRMGASRRCKSASRSMCSMRMCKMRGSCWSSKMMGSGRSRA 17692
Qy 1028 agatggcgagcagcattgctgcgaagaaggaatccgcgcagcagcacaatttgcgcag 1087
Db 17693 SSKCKGRCMRMRKRSKSKYKRGKRRSMTKSGSKGSKCWRSGSMSTSCYYYSAS 17752
Qy 1088 cggcagcagccaataccgctcccttaccattcccggaatattccgltcaaatctggagc 1147
Db 17753 SCMMMSKSCMCCMMKRRKACCYSMSCTSMYRCMGWGMKSYSTCCTTCKKCTGYRK 17812
Qy 1148 agcgltcggaagaagaacaatccatccctcagccgctgcgcgcagcagcaaatc 1207
Db 17813 SKTYWMMWYMTSKMYMRAACMYCYMSTRMCMMSMCAKCKMCMCTMKAGATGACA 17872
Qy 1208 tcaactgycagaccac 1225
Db 17873 GTAGAGTAGATAGAAC 17890

RESULT 41
AF429315/c 125020 bp DNA linear PRI 18-JAN-2002
LOCUS AF429315
DEFINITION Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 125020)
AUTHORS Holmes, S.E., O'Hearn, E., Rosenblatt, A., Callahan, C., Hwang, H.S.,
Ingersoll, A., Ashworth, R.G., Fleisher, A., Stevanin, G., Brice, A.,
Potter, N.T., Ross, C.A. and Margolis, R.L.
A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
Nat. Genet. 29 (4), 377-378 (2001)

TITLE
JOURNAL
MEDLINE 11694876
PUBMED 21583737
REFERENCE 2 (bases 1 to 125020)
AUTHORS Holmes, S.E., Ingersoll, A., Ashworth, R.G., Ross, C.A. and Margolis, R.L.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
FEATURES
SOURCE
1..125020
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q24.3: between DIS520 and WI-12410"
/note="Isolated from a patient with Huntington's
Disease-Like 2 (HDL2)"
complement(35361..35746)
/rpt_type=tandem
/rpt_unit=ctg
complement(<36507..>36887)
/gene="JPH3"
/product="Junctophilin 3"
complement(<36507..>36887)
/gene="JPH3"

CDS
/note="JPH3"
complement(<36507..>36887)
/gene="JPH3"
/note="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/product="Junctophilin 3"
/protein_id="AAL40941.1"
/db_xref="GI:17646245"
/translation="MSGGRFNPDDGSGYCGGMDCKAHGVCCTGPKGGEYTSMS
HGFEVLGVTPSGNTYOGTWAQCKRRLGLESKKRYVGEWTHGKRGVRECAQ
NKAKEGTWSNGLDDEGTETYSIDG"
BASE COUNT 29056 a 32731 c 30696 g 28283 t 4254 others
ORIGIN

Query Match 2.68; Score 48.4; DB 9; Length 125020;
Best Local Similarity 12.68; Pred. No. 1.8;
Matches 124; Conservative 393; Mismatches 457; Indels 7; Gaps 3;

Qy 837 agatattgagcaactcaagaactatgcccagcagcagccatccgcatgtggcagtcacaa 896
Db 52150 RGRGMRKTYRGMWRRSMCMSCCYCMCKCACMSMAKCMCKSCYMYSMYRGSMSTGC 52091
Qy 897 ccccaatgcccagcaagcagatagaagccgtaagcaatattttagagccatcccat 956
Db 52090 CMCKYKRSRSMSCYKGSWSGYWASMSYKSCASCWGSMTYSRSGCWATSRYCKMYGS 52031
Qy 957 caaaggattggagctgctcggggaaatacagccttggtggcgagcagcagcattcgt 1016
Db 52030 KTRSGTGKRSACMCKWCMWSSNSR--SMCMYAMCMWCMWYRYRSRRRYSCYRS 51974
Qy 1017 caagcgtcgcagatggtgcgcatgcgcatcccggaagaaggaatccgcgtcagcagac 1076
Db 51973 KCYRSRAKSMAMMGSMMSYVMMCMCKGCRGAGSAARSYKKAAMSKGMRGCG 51914
Qy 1077 tttagcagatgagcagcagcagcaataccgctcccttaccattcccggaatattcgtc 1136
Db 51913 STYYGCKMGKSGSKMYTSGMCAATGAKMSWSMCMGMRGMSGGSYYRGWCMWCMYSYW 51854
Qy 1137 aaacttgagcagcgttcagcgaagaagaacaatccctcctcagcagcgtccgcgtcaaa 1196
Db 51853 CASYSRMRGSAMSTWAGCTYRKGSW---GGMKMYKSRMRGSMRSMKSKSMKSMR 51797
Qy 1197 cggcgaatattcaaacgycgagcagcagcagcagccacccaggaagcagcgtttagcgg 1256
Db 51796 KGRSRASRMRKSYSCASRRKSCWRMRMGCSYMTGRSSMGTMSKRGSRPSMSKAGTYRS 51737
Qy 1257 taagggttccgaatttttagaagcagcgtgaataatgatgacgaagctcgatatcaaga 1316
Db 51736 ARRAAMYRSCSGKRRRRGKSSKSGGGRGCKWMAWSMTYSKRCACSTGRRRGCTGSSCW 51677
Qy 1317 attatcgggggcggtatacctaagaagcgttglttgatgagcaacaggaatggga 1376
Db 51676 TKRRGWGTRRAAAMCAACCMWYRCTGCMKATYATYTKSMKSWTGTWRMRGTRRA 51617
Qy 1377 ggttataggaagccttaaatltgacaactcgtgagcagcagtgaggaataatgttcaaga 1436
Db 51616 MWWMYMWWKYYARMKCAKTRBMMAWSKYSKSYARCCSMMSYAKKYYKRSASRCKWGM 51557
Qy 1437 aacgagaagaagagtgatcagatgagcgtttaaagccatctgcagcaggaat-gggaa 1495
Db 51556 AKSCMRRYMMYRWGGAGCCGAGTGTRGSWSACCKMSKSSKAYATAYSTGA 51497
Qy 1496 ataacaagggtagatttaattatcattatagtggttgatatacaataagaagagcag 1555
Db 51496 MCMGCTCYMAKCTTRMMATACAMRRRAWMSCMKCMYTSKSKMGRRRYCWSNAR 51437
Qy 1556 taacagagagcagatgcttaaccgctggtgagtgatgacgagtaacaaacctggcac 1615
Db 51436 YVAGSYMSKCRMRGWSYKGRSRSGSMWRTRGYSYARCYMSKSGAGACKGAGSYAK 51377
Qy 1616 ctgataaacatggttttatacagcagcagtggaatataaagccctggaagttggg 1675

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 29, 2002, 17:16:22 ; Search time 27321.5 Seconds

(without alignments)
1029,420 Million cell updates/sec

Title: US-09-303-518D-125

Perfect score: 1344

Sequence: 1 atgattaaatacaaaaaagg.....ccatcgagaagaagcctga 1344

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 segs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : GenEmbl:*

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Match Length DB ID Description

1	1344	100.0	1344	6	A96092
2	1344	100.0	12273	1	AE002412
3	1344	100.0	349980	6	AX044029
4	1244.8	92.6	311321	1	NMA352491
5	1229.4	91.5	1344	6	A96094
6	1212.8	90.2	1344	6	A96098
7	560	41.7	10563	1	AE006171
8	556.6	41.4	1558	1	AFU24492
9	524.6	39.0	13299	1	U32702
10	505.6	37.6	6048	1	AF165980
11	489.6	36.4	5916	1	AF117331
12	489.6	36.4	13382	1	AE004300
13	477.2	35.5	16903	1	AE004724
14	475.2	35.4	5078	1	VANORBL
15	473.6	35.2	474	6	A96090
16	473.6	35.2	6908	1	AE008030
17	436.4	24.3	96109	6	AF144076
18	413.6	8.5	1356	1	AF144076
19	410.6	7.5	4926	1	HTU20229
20	47.4	3.5	125020	9	AE429315
21	46	3.4	36200	1	BACGNTA
22	46	3.4	21534	1	BSUB0021
23	41.4	3.1	109730	2	OSJN00193
24	41.4	3.1	142805	2	OSJN00146
25	40.2	3.0	151329	2	AP003456
26	39.8	3.0	216050	1	AL646076
27	39.2	2.9	24268	1	AB070947
28	39.2	2.9	187948	9	AE003402
29	38.2	2.8	37245	1	SC5F2A
30	38.2	2.8	38734	1	SC10A5
31	38	2.8	975	1	AF103943
32	37.8	2.8	1694	8	ATY14325
33	37.8	2.8	11351	1	AE004619
34	37.8	2.8	11660	1	AE004956
35	37.8	2.8	103573	8	AC005499
36	37.6	2.8	10944	1	AE002563
37	37.6	2.8	17325	6	AX044035
38	37.6	2.8	198050	1	AL646061
39	37.6	2.8	340806	1	NNA122491
40	37.4	2.8	4499	8	AF324244
41	37.4	2.8	7365	1	SM5SVDEOX
42	37.4	2.8	39250	1	SV1011500
43	37.4	2.8	125020	9	AF429315
44	37.4	2.8	154330	2	AC104429
45	37.2	2.8	3867	1	AV033407
46	37.2	2.8	4130	8	AF190427
47	37	2.8	2557	3	AV071018
48	37	2.8	5021	1	AX024382
49	37	2.8	5021	6	AX024275
50	37	2.8	37750	1	MICB1913
51	37	2.8	89791	3	AC005268
52	37	2.8	101085	2	AC018341
53	37	2.8	164035	3	AC007888
54	37	2.8	213050	1	AL646079
55	37	2.8	269203	1	MEPRN10
56	37	2.8	305020	3	AE003452
57	36.8	2.7	10500	1	AE005062
58	36.8	2.7	17589	1	AE004862
59	36.4	2.7	1725	6	AX073998
60	36.4	2.7	2076	6	AX073995
61	36.4	2.7	2109	6	AX350975
62	36.4	2.7	10486	1	AE001656
63	36.4	2.7	15680	1	AE002164
64	36.4	2.7	33801	1	SCC77
65	36.4	2.7	38634	1	SC2665
66	36.4	2.7	300550	1	AP002547
67	36.2	2.7	4623	1	AF419157
68	36	2.7	760	4	AY011741
69	36	2.7	975	4	AY011738
70	36	2.7	975	4	AY011739
71	36	2.7	975	4	AY011740
72	36	2.7	11526	1	AF228498
73	36	2.7	349116	1	AP003003

A96092	Sequence	12
AE002412	Neisseria	
AX044029	Sequence	
AL162754	Neisseria	
A96094	Sequence	12
A96098	Sequence	13
AE006171	Pasteurel	
U24492	Actinobacil	
U32702	Haemophilus	
AF165980	Vibrio ha	
AF117331	Vibrio ch	
AE004300	Vibrio ch	
AE004724	Pseudomon	
Z37111	V. alginolyt	
A96090	Sequence	12
AB008030	Vibrio al	
AX067460	Sequence	
AF144076	Porphyrin	
U20229	Haemophilus	
AF429315	Hom sapi	
D78193	Bacillus su	
Z99124	Bacillus su	
AL662992	Oryza sat	
AL662935	Oryza sat	
AP003456	Oryza sat	
AL646076	Ralstonia	
AB070947	Streptomy	
AP003402	Hom sapi	
AL021529	Streptomy	
AF103943	Streptomy	
Y14325	Arabidopsis	
AE004619	Pseudomon	
AE004956	Pseudomon	
AC005499	Arabidops	
AE002563	Neisseria	
AX044035	Sequence	
AL646061	Ralstonia	
AL162754	Neisseria	
AF324244	Phaseolus	
L37334	Streptomyces	
AF011500	Streptomy	
AF429315	Hom sapi	
AC104429	Oryza sat	
AY033407	Myxococcu	
AF190427	Colletoct	
AY071018	Drosophil	
AX024382	Sequence	
AX024275	Sequence	
AL022118	Myobacte	
AC005268	Drosophil	
AC018341	Drosophil	
AC007888	Drosophil	
AL646079	Ralstonia	
AL583926	Myobacte	
AE003452	Drosophil	
AE005062	Halobacte	
AE004862	Pseudomon	
AX073998	Sequence	
AX073995	Sequence	
AX350975	Sequence	
AE001656	Chlamydia	
AE002164	Chlamydia	
SCC77	Streptomy	
AL136503	Streptomy	
AP002547	Chlamydia	
AF419157	Aeromonas	
AY011741	Canis fam	
AY011738	Felis cat	
AY011739	Leopardus	
AY011740	Panthera	
AF228498	Escherich	
AP003003	Mesorhizo	

ALIGNMENTS

338

CH

```
/organism="unidentified"  
/db_xref="taxon:32644"
```

111

0000

26

AAG

99

Db 1261 TTGTCACGCTCTGTCGCCGCGCAATACGATACGCGCCGCTGTGCGCAAGTCTG 1320
Oy 1321 gaaccattgagaaggaagctga 1344
Db 1321 GAAACCATGTGAGAGGAGGCTGA 1344

RESULT 2
AE002412/c
LOCUS
DEFINITION Neisseria meningitidis serogroup B strain MC58 section 54 of 206 of the complete genome.
ACCESSION AE002412 AE002098
VERSION AE002412.1 GI:7225783
KEYWORDS
SOURCE
ORGANISM Neisseria meningitidis MC58.
Neisseria meningitidis MC58.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;

REFERENCE
AUTHORS
1 (bases 1 to 12273)
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwinn, M.L., Deboy, R., Peterson, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E., Clifton, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H., Qin, H., Vamathevan, J., Gill, J., Scariato, V., Maignan, V., Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R., Rappapoli, R. and Venter, J.C.
Complete genome sequence of Neisseria meningitidis serogroup B strain MC58
Science 287 (5459), 1809-1815 (2000)

JOURNAL
MEDLINE
PUBMED
10710307
2 (bases 1 to 12273)
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwinn, M.L., Deboy, R., Peterson, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E., Clifton, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H., Qin, H., Vamathevan, J., Gill, J., Scariato, V., Maignan, V., Rappapoli, R. and Venter, J.C.
Direct Submission
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

TITLE
JOURNAL
FEATURES
SOURCE
Location/Qualifiers
1. 12273
/organism="Neisseria meningitidis MC58"
/strain="MC58"
/db_xref="taxon:122586"
/note="serogroup: B"
97. 435
/gene="NMB0557"
97. 435
/gene="NMB0557"
97. 435
/note="conserved hypothetical protein; identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA040985.1"
/db_xref="GI:7225784"
/translation="MSDESLIFDSCCAVADLIAENNPDILKRVFVNGGSCGFO YGTFEDIKNDDEFEIKNGLVLPVPMSTQYLVGAIDYTESLQSGQVYIRPNMAYT TCGGGSSEFSY"
595. 1059
/gene="NMB0558"
595. 1059
/gene="NMB0558"
/note="hypothetical protein; identified by Glimmer2; putative"

gene
CDS
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA040986.1"
/db_xref="GI:7225785"
/translation="MIEMSKDYRNDLADYVVSYPQVDRGLIRECLKENLGEKEAGEL IESLDSKPOVLVEEKECTWAKREELADYFSYGLDITRRVMELETVPVPEEGEGEG ADGEMPEYLELHGREDDISAPQPEPSRNKILVPLGLIAFLGYLLKIF"
1094. 2605
/gene="NMB0559"
1094. 2605
/gene="NMB0559"
/note="similar to GB:M87049 SP:P27853 PID:148233 GB:U00096 PID:2367309 percent identity: 74.73; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ubiquinone biosynthesis protein AarF"
/protein_id="AA040987.1"
/db_xref="GI:7225786"
/translation="MKWLKRLVIVGFYRYRLAGLCAISLMSGMICALKMPQSSK LKNEPPAVRLIALESIGPIFIKFGVLTSPDILPHDYAVEALAKLDQVPEPARLS RQIKSLSGOSIEKLIAEFETEPASIASIAOVHRAKLHSGQVAVKYLPLLELVIG DLSIMRFAGWYERLEFADGKRLKPEVVAEEDKYLHDELDMREANASOLGRNONS DMLIVKVEYVCTSDVLTIMMOCPEPSDIAKKADSIDKLADYGVLEFFQVNR DCFIADHPRNIIYAADNRITADREGVITLTQDKRYLAINTLAFPNBRYRVATA HIESWVPAQDRAEELAEAVAVCVSPVNFPSISITSELVLMKLEFVSRRNVEIQO LVLLQKTLNTEIGLEIROPDLDMKTKAPPLVKNMGQVGPALMKNLKEAPDMQO IIPSLPKRISALIDENRQEMRDYAIHLVKKVOQOSIMLAVIAVALLILLK"
/complement(2962. 3780)
/gene="NMB0560"
/complement(2962. 3780)
/note="similar to GB:I42023 SP:P43886 PID:1005429 PID:1220683 PID:1204657 percent identity: 72.97; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="serine acetyltransferase"
/protein_id="AA040988.1"
/db_xref="GI:7225787"
/translation="MKKHDLNTGTGFDLMTITREETAASVSEPMPLASTIRHOTVRRHS LGSVLAYHSSRLSPIMDVRLFEITVOALGSDIQGKCEADLAKIETEDPCDEL SILLIFKGFPAIQAHRIHRIYLDGRITATVFLQNRSEVFGVDIHPAEPFGDML DHATCFVAGETAVLGNISILHGYTLGSGEGSDRHKKIDGVMIGAMASILGNIRI GSNATIGAGSVVSDVPISITVGVPAKPARSIRKTPSADMDONIOETIDPMI"
3964. 4542
/gene="NMB0561"
3964. 4542
/gene="NMB0561"
/note="similar to GP:3777486 percent identity: 61.03; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="grp94 protein"
/protein_id="AA040989.1"
/db_xref="GI:7225788"
/translation="MSBOTCOONSEFENVEVETVETVTCNAGQVOBOAAAEAYED LOARITAELOQKDEQLALANBOMLRRHQEILADYHKFPGQRFAYEMLVKDTLEA ALLDOSGNFDALKMGVOMTLELQAFATQIKETNPRAQKLDPNIHQAQAVASQ EPNTVGVWVKKQYTLSDRLVLRPAMVTAOKBA"
/complement(4819. 5034)
/gene="NMB0562"
/complement(4819. 5034)
/gene="NMB0562"
/note="conserved hypothetical protein; identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA040990.1"
/db_xref="GI:7225789"
/translation="MKTLTLTGIFLVYIIGMAYGIRSKRSINGSCGCTIALCKMKM

CDCTPCTDTRLOKKLDEKQAGIRVDR"
complement(5060. .6115)
/gene="NMB0563"
complement(5060. .6115)
/note="similar to GB:142023 SP:P44550 PID:1003244
PID:122087 PID:1204428 percent identity: 74.70;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="thiamine biosynthesis lipoprotein AbpE"
/protein_id="AAF40991.1"
/db_xref="GI:7225790"
/translation="MSETRLPNFIYVILFALGFILNCSQEQTAQTVTLIGFTMGTT
YVAVKYSNNRDKLPSPAEIQKRIIDALKEVNRQMSYQPSSEIRSNQHTAGKPLRIS
SFAHAYTAVALRLNLTGALDVTGAPLVNMGFGDKSVTRSPSPQIKOASVTGI
DKILKQGDYASISLSTHPRAYLDLSSIAKGFVDVAGLEKEGYIONLVEIGSELI
GKKNRGEPMWRIEOPNIVOGNTQIIVPLNRSLATSGDYRIFHVDKNGKRLSHI
IDNNRPRISHNLASISYVADSMTADGISTGFIYGEETALKLAREKLAFLIYLRD
KGGRTAMSEPEKILR"
complement(6268. .7485)
/gene="NMB0564"
complement(6268. .7485)
/note="similar to GB:L42023 PID:1003242 PID:1222086
PID:1204427 PID:1573127 percent identity: 92.82;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="Na(+)-translocating NMDH-quinone reductase,
subunit F"
/protein_id="AAF40992.1"
/db_xref="GI:7225791"
/translation="MEIILGIWFTYIVLVALLMLFAKSKLIVSEGDITIKVGEKEL
TWPAGSKILGALANEGIFIPISAGGSGGCGRVYKSGGDLITPELISHISRRRE
GCRISQVYVKTDMDEVEPEVGVKMECTVYSNDRKATFKEILALPEGEVEFR
AGGYTOIEAPPHATKDDIREHEDMDKTNLMQVSKVDEPILIRATSMASYPEEK
GIMLVNRIATPPRPVADAPGOMSSYIWSLRPGDVTISGPEGEFAADTAEWFI
GGAGAAAPKRSHPDLKRLNRSKRITFEYGARSKREMFVEEDDLAEPFETWVH
ALSDPLPEDNMDGYTGFIHNVYENHLKHEAPEDEFCYCPINQSVIKMLKDLG
VEDENILDDPEG"
complement(7499. .8092)
/gene="NMB0565"
complement(7499. .8092)
/note="similar to PID:1573126 percent identity: 91.92;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="Na(+)-translocating NMDH-quinone reductase,
subunit E"
/protein_id="AAF40993.1"
/db_xref="GI:7225792"
/translation="MEHYLSLFIKSVFIENMALISFLGMCPTFLAVSKVYATGILGA
VTFVGLSVNQLVYSLKDGAIAGVDLTFELKFTFGVIAALVLIEMFLDKFVP
ALYNLGIYLPITVNCALFCAVSEMAQREYNGESVYVFGAGLGMALAIVALGIT
EKMKYSDAPKKGKLGITFIAGLAMAMAFMSFSIQLV"
complement(8096. .8722)
/gene="NMB0566"
complement(8096. .8722)
/note="similar to GB:L42023 SP:P43958 PID:1003237
PID:1222083 PID:1204425 percent identity: 89.42;
Query Match 100.0%; Score 1344; DB 1; Length 12273;
Best local similarity 100.0%; Pred. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 12009 TACGACGCCCCGGCCATTACCGAAGTGCCTGCTTGGCCGAAGAAATGTGCCGGATATCCG 11950
|||
QY 121 ccctcgaatgaagtcgaagagcgatgcgctcaaaaaagccaagtgtcttgaagac 180
|||
DB 11949 CCCGCGATGAAGACCAAGCAAGCGAGTCCCTCAAAAAAGCCCAAGTGTCTTGAAGAC 11890
|||
QY 181 aaaaagaatccggcggtgtgttactgtgcgcgcgcttcgaagcaaatcgcgcgattcac 240
|||
DB 11889 AAAAAGAAATCCGGCGGTGTGTACTACGCGCCGCTTCAGGCAAAATCGCGGATTCCAC 11830
|||
QY 241 cgtgagcaaaagcgagcttcaagtcagtcagtcagtcagtcagtcagtcagtcagtcagtc 300
|||
DB 11829 CGTGGCAAAAAGCGCGTACTTCAATGATGATGATGATGATGATGATGATGATGATGATG 11770
|||
QY 301 gagttgaacgtacagcaactgaagcgctgcaaaactgaagcgagcgagcgagcgagcgagcg 360
|||
DB 11769 CAGTTTGAACGCTACGACGACTGAAGCGCTGCAAACTTAAGCGGCGAAGTGGCGCGC 11710
|||
QY 361 aacctgattccatccggtttgtgactgagcgtgcgacccgctccgttcagcaaatctct 420
|||
DB 11709 AACCTGATCCAAATCCGGTGTGTGACATGCGTCCGCAACCGCTCGTTCACCAAAATTCCT 11650
|||
QY 421 gccgtcagatccgagccgttcgcacatctcgtcaatgcagtcgagcaaatccgctgct 480
|||
DB 11649 GCGGTGATGCCAGCGCGTTCGCGCATTTGCTCAATGCAATGGACACCAATCCGCTGCT 11590
|||
QY 481 gcgagcccaagtcgacatcatcaaaagacgcgcgagagattcaacgcgagcctgttgta 540
|||
DB 11589 GCGGACCCCTACGGTCAATTAACAAAGAACCGCGAGATTTCAACGCGGCTCTGTGTGTA 11530
|||
QY 541 ttgagccgtttgacgcgaagcaaatccatctgtttgaagcaactgagcgagcgagcg 600
|||
DB 11529 TTGAGCCGTTTGAACCGACCAACCAATTCATGTTGTAAGGCACTGGCGGACGTTGCCG 11470
|||
QY 601 tctgaataatgctcgaacatcgaaacacatgaattcggcgcgcgcgcgcgcgcgcgcgcgcgc 660
|||
DB 11469 TCTGAATAATCTCTCCAAATCGAAGCAACATGAATTCGCGCGCGCGCATCTCGCGGTTTG 11410
|||
QY 661 agtgcgacgacatctcatcgagcgcgctgcgcgcgagtaaaacgcgtgtgacatc 720
|||
DB 11409 AGTGGCAGGACATTCATTTCATCGAGCGCGTGGCGCGCAATTAACCGTGTGACCATTC 11350
|||
QY 721 aattatcaagaatgaatccatctgagcgctgtgttgcaagcgagcgctgtaacccgag 780
|||
DB 11349 AATTATCAAGATGTAATACCATTTGGCGGTTGTTGCAACAGCGCGTGTGAACACCGAG 11290
|||
QY 781 cgcgtgattgcctcagtggttctcaagtcacaacacgcgcgcctctgtgtaacgctttg 840
|||
DB 11289 CGCGTATTCCTTAGTGGTTCTCAAGTCAACAAACGCGCGCTTGTGCTACGTTTGG 11230
|||
QY 841 ggtgcgaagatcgcgaatctcgcgagcgagctgtgtgaacacagacacgcgctgatt 900
|||
DB 11229 GGTGCGAAGATATCGCAAAATTAATCTGGCGGGAATTTGTTGACACAAACCCCGTGAAT 11170
|||
QY 901 tccggttcgatttgaacgcgcgcatlacaaagcgcgacgaatatttggagcctac 960
|||
DB 11169 TCCGTTTCGTTATGAAGCGCGGATTTACCAAGCGCGCGCATATTATTGGACGCTAC 11110
|||
QY 961 cacaatcagattccgltatcgaagaagcgcgagcaagaagctgttcgctgtggttgcg 1020
|||
DB 11109 CACAAATCAGATTTCCGTTATCGAAGAAAGCGCGCAAGAAAGCTGTTCGCTGCTGGTTGGG 11050
|||
QY 1021 cgcgacgcgagacaatctacatccacgcgctacacacccctcgacattctcgtgaacaa 1080
|||
DB 11049 CCGCAGCCGACAAATATCTCATACGCGTACACACCTCGGCGATTCTTGAAAAACAAA 10990
|||
QY 1081 cttctcaagltcaacacagcgctcaacgycgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1140
|||
DB 10989 CTTCTCAAGTTCAACACAGCGCGTCAACAGCGCGGACCGCGCAATGCTGCGATGTTGACT 10930
|||
QY 1141 taccgacgcggtgagtccttgatlatccgtccacaccccttcttgagatttaacgctc 1200
|||

Db 10929 TACGAGCGGCTGATGCCCTTGATATCCCTCCACCCCTGCTTTGGCGATTATATGTC 10870
Qy 1201 ggcgataccgaagcagcgacgagcattgggtgtgcttgaaatgagcaagaagacctgcct 1260
Db 10869 GCGGATACCGACAGCGCGAGCATGGGTGCTTGGAATGGACGAAGAAGACCTCGCT 10810
Qy 1261 ttgtgagcttgcgtgcgcggcgcaaatagcaatagccgcgtgtgttgcaagaagtgtg 1320
Db 10809 TTGTGAGCTTGGCTTGCCTCCGGCAATATGCAATACGCGCCGCTGTGTGCGCAAAAGTGTG 10750
Qy 1321 gaaccatcgaagaagaagcgtga 1344
Db 10749 GAACCATTTGAGAGAAGAGCTGA 10726

RESULT 3
AX044029/c 349980 bp DNA linear PAT 24-NOV-2000
LOCUS Sequence 108 from Patent WO0066791.
DEFINITION AX044029
ACCESSION AX044029 GI:11342913
VERSION AX044029.1 GI:11342913
KEYWORDS
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis
Bacteria: Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria
REFERENCE 1 (bases 1 to 349980)
AUTHORS Pizza, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J. C.,
Masiyana, V., Galeotti, C., Mora, M., Ratti, G., Scariselli, M.,
Scazzato, V., Rappunli, R., Frazer, C. M. and Grandi, G.
TITLE Neisseria genomic sequences and methods of their use
JOURNAL Patent: WO 0066791-A 108 09-NOV-2000;
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
source location/Qualifiers
1..349980
/organism="Neisseria meningitidis"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces. -seq 1: 1 to
349980 349980 bases--seq 108: 300001 to 649980 349980
bases--seq 109: 600001 to 949980 349980 bases--seq 110:
900001 to 1249980 349980 bases--seq 111: 120001 to 1549980
349980 bases--seq 112: 150001 to 1849980 349980 bases--seq
113: 1800001 to 2149980 349980 bases--seq 114: 2100001 to
2272325 172325 bases"

BASE COUNT 84410 a 84863 c 94187 g 86520 t
ORIGIN

Query Match 100.0%; Score 1344; DB 6; Length 349980;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgattaaatcaaaaaagcttaacctgcccacgcgcggcgacgcggagcgaacgcgtt 60
Db 297909 ATGATTAAATCAAAAAAGGCTTAACCTGCCATCGCGGCGAGACCGGAGACCGCTT 297850
Qy 61 tacgaacgcccgcgcattacgaagtcgctgtgttgcggaagaatatccggtatgcgc 120
Db 297849 TACGACGGCCGGCCATTACCGAAGCTCGCGTGTGGCGAAGATATCCCGGTATGCC 297790
Qy 121 ccctcgaatgaagtcgaagaagcgatgcgtcaaaaaagcgcaagtgtcttgaagac 180
Db 297789 CCTTCGATGAAGATCAAGAGAGGATGCGGTCAAAAAAGCGCAAGTCTGTTGAAGAC 297730
Qy 181 aaaaaaatccggcgctgtgttctacitcgccggttcaaggcaaatccgcgattcac 240
Db 297729 AAAAAGAAATCCGGCGGTGTGTACTCGCGCGCTCAAGCAAAATCCCGGATTTAC 297670
Qy 241 cgtgacgaagaagcgctactcaatcagtcagtcgtgttgccgttgaagaagcgaacgaatc 300
Db 297669 CGTGCGCAAAAGCGCGTCTTACATGATCGTGTGATTCGCCGTTGAAGGCAACGAAATC 297610
Qy 301 gatttgaaacgtacacactgaagcgtctgcaaaccttaagcgcgaagaagtgcgcgc 360

Db 297609 GAGTTGAACGGCTACCGCACCTGTAAGCGCTGGCAAACTTAAGCGGCGAAGATGCGCGC 297550
Qy 361 aacctgatccaatccggttctgtgagctgcgtgcgaacccgttcctgttcaacaaattcc 420
Db 297549 AACCTGATCCAAATCCCGTTGTGTGAGCTGCGCTCGCACCGCTCGTTCACCAAAATTCCT 297490
Qy 421 ggcgttcgatgcgcgaagcgtttgcacatcttcgtaatgcgaatgcgaatgcgaatgcgt 480
Db 297489 GCGGTGATGCGAAGCGCGTTCGCCATCTTGTGTAATGCGATGAGACACCATCCCGTGGCT 297430
Qy 481 ggcgaacctacggttcaatcaagaagccgcggaagatttcaaacgcggtgttgta 540
Db 297429 GCGGACCTTACGGTCAATTATCAAAAGACCGCGGAGGATTTCAAAACCGCGCTGTTGTA 297370
Qy 541 ttgagccgttttgcaacgcaagcgaatccatgatttgtaagcagctgcgcgaagcgtccg 600
Db 297369 TTGAGCGCTTTGACCGCAACCAAAATCAGTTGTGTAAGCAAGCTGCGCGAGAGCTGCG 297310
Qy 601 tctgaaaatgcctgcacacatcgaaacacatgaaatcgcgcgcgcgcacatccctgcggttg 660
Db 297309 TCTGAAAATGCTTCCCAATCGAAGACATGAAATTCGCGCGCGCATTCCTGCGGTTTG 297250
Qy 661 agtgcacgcacatcattatcgcagcgcgtgcgcgcgaatcaaacgcgtgtgacatc 720
Db 297249 AGTGCGACGACATTCATTCATTCATCGAGCCGCTCGGCGCAATTAACCGTGTGACCATC 297190
Qy 721 aattatcaagatgtaattaccattggtgcgtgtgttttgcacagccgcttgaacacgag 780
Db 297189 AATTATCAAAATATTAATTAATCAATGCGCGTGTGTGTGCAACAGCCGCTGTCAACACGAG 297130
Qy 781 cgcgtatgcctcagtggttctcgaagtcacaacacgcgcctcttgcgtacggtttg 840
Db 297129 CGGCTATTCCTCAGTGTGTTCTCAAGTCAACAAACCGCGCTCTTGCCTGACGTTTG 297070
Qy 841 ggtgcgaagatcgcgaatcattctgcgcgcgtgtgttgacacagacacgcgtgtgatt 900
Db 297069 GGTGCGAAATATTCGCAAAATTAACGCGCGCAATTTGTTGCAAGAACCGCGCTGATT 297010
Qy 901 tccggttcggttatgaaacgcgcgcattacacaaagcgcgacagattatitggagcgtac 960
Db 297009 TCCGCTTCGCTATTTGAACGCGCGCATTAACAAAGCGCGCACGATTTATTTGGAGCTAC 296950
Qy 961 cacaatcagaattccggtatcgaagaagccgcgaagaagcgtgttcgcgtggtgcg 1020
Db 296949 CACAATCAGATTTCCGTTATCGAAGAGCGCGCAAAAGAGCTGTGCGCTGGGTTGCG 296890
Qy 1021 ccgcagccgcgaacaaatcctcatcagcgtctacaacccctgcgcattcttgaaaaaacaa 1080
Db 296889 CCGCAGCGCGAACAATATCTCCATCAGCGGTACCAACCTCGGSCCATTTCTGAAAAACAA 296830
Qy 1081 ctcttcaagttcaacacacagccgtcaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1140
Db 296829 CTCTTCAAGTTCAACACAGCGGTCAACGCGCGCGCACGCGCATGTGTGCTACT 296770
Qy 1141 tacgagcgcgtgatgccttggatataccgtcccaacctgtcttgcgcgattaaatgc 1200
Db 296769 TACGAGCGCGGATGCTTCGATATCTCGCCACCTCGCTTTTGGCGGATTTAATTCG 296710
Qy 1201 ggcgatccgaacgc 1260
Db 296709 GGCATACCGGACAGCGCGGAGGCAATGGGTGTGTTGGATTGGACGAAGAACCTCGCT 296650
Qy 1261 ttgtgagcttgcgtcctccgcgcgaataatagcgcgcgttltgcgcaaatgtgtc 1320
Db 296649 TTGTGAGCTTGCCTGCGCGGCAAAATAGATATGCGCGCTGTGGCAAAAGTGTG 296590
Qy 1321 gaaccatcgaagaagaagcgtga 1344
Db 296589 GAACCATTTGAGAGAAGAGCTGA 296566

RESULT 4

NMA322491/c 311321 bp DNA linear BCF 04-DEC-2000
LOCUS NMA322491
DEFINITION Neisseria meningitidis serogroup A strain Z2491 complete genome;
ACCESSION AL163754
VERSION AL163754.2 GI:7379424
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
1 (bases 1 to 311321)
Parkhill, J., Achtman, M., James, K. D., Bentley, S. D., Churcher, C., Klee, S. R., Morelli, G., Basham, D., Brown, D., Chillingworth, T., Davies, R. M., Davis, P., Devlin, K., Feltwell, T., Hamlin, N., Holroyd, S., Jagels, K., Leather, S., Moule, S., Mungall, K., Quail, M. A., Rajandream, M. A., Rutherford, K. M., Simmonds, M., Skelton, J., Whitehead, S., Spratt, B. G. and Barrall, B. G.
Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
Nature 404 (6777), 502-506 (2000)
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Details of N. meningitidis sequencing at the Sanger Centre are available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
Location/Qualifiers
1. 311321
/organism="Neisseria meningitidis Z2491"
/strain="Z2491"
/db_xref="taxon:122587"
/feature="serogroup: A"
115. 120
129. 587
139. 587
/feature="NMA0698"
/feature="NMA0698", possible periplasmic protein, len: 152 aa; unknown, contains a probable N-terminal signal sequence and lies within a region of unusually low GC content
/codon_start=1
/transl_table=11
/product="putative periplasmic protein"
/protein_id="CAB83985.1"
/db_xref="GI:7379425"
/db_xref="SPTREMBL:Q9J0Q2"
/transl_table="MLTKLILFLFVLAINILFFPSSDIESPNGYQFXYVD KGPANYILVMKDGNEGNFDITSGLEVYKEDDNTYFSTIDGQFASDSCYKPEI LYKGLINKNHNIIINSMKNNFLSDEKIMKGRNWLADPRKNCNIQTLTD"
742. 1077
/feature="NMA0699"
/feature="NMA0699", pseudogene, probable ABC transporter protein, len: 336 bp; similar to C-termini of many ABC transporters e.g. TR:Q46973 (EMBL:U47048), mtfB, Escherichia coli microcin transport protein (707 aa), fasta scores: E(): 2.6e-19, 53.2% identity in 111 aa overlap. Contains PS00211 ABC transporters family signature. NMA0686 may be the remainder of this pseudogene
/codon_start=1
/pseudo
/transl_table=11
/product="putative ABC transporter protein (pseudogene)"
868. 912
/feature="NMA0699"

/note="PS00211 ABC transporters family signature"
complement(1101..2327)
/feature="NMA0700"
complement(1101..2327)
/feature="NMA0700"
/feature="NMA0700", possible ribonuclease BN, len: 408 aa; similar to SW:RBN_ECOLI (EMBL:U19201), rbn, Escherichia coli ribonuclease BN (Ec 3.1.-.-), fasta scores: E(): 2.3e-28, 34.3% identity in 268 aa overlap. Longer than rbn at the C-terminus. Also similar to TR:085449 (EMBL:AF067083) Vitreoscilla sp. hypothetical protein (376 aa), fasta scores: E(): 0.45e-0% identity in 333 aa overlap. Contains a region similar to NMA0677, fasta scores: E(): 4.3e-08, 64.4% identity in 45 aa overlap. Contains hydrophobic, possible membrane-spanning regions, as does rbn
/codon_start=1
/transl_table=11
/product="putative ribonuclease BN"
/protein_id="CAB83987.1"
/db_xref="GI:7379426"
/db_xref="SPTREMBL:Q9UV79"
/transl_table="MTFLQRLGLADNKCICAFAMFVRRDERVPOAASMTFTLL ALVPLVVMVAVASIFPVEDRSDSEFVNOTIYOGADMYDYINAFEOANRLTA IGSVMALVYTSMLRLTIDTENTRIRVNSORPMQOFLYVWALITFGPISLIGISPM VGSYODALASGAPQWGSALRAATLTETLLKCLYRVRPREVPARQAFVALATA FCLERTASLPTWYKGNFPGYSITAFPAVPPFLMLNLTMLVLGAVTSLSTWQ GEATRGKDSGRFDVDLKLILLDAAKGEKALPVEFRHITNMGYDELGLEKIA RHGIYISGRQGWVLTGTADSIENLFLFVYRPLPVERDHVNOAVDAVWMPCLQTLN MTLAEFPAQARKKQOOS"
1376. 1385
/feature="Core DNA uptake sequence: gccgcgtctgaa"
/label=DUS
complement(2417..2426)
/feature="Core DNA uptake sequence: gccgcgtctgaa"
/label=DUS
2444. 2453
/feature="Core DNA uptake sequence: gccgcgtctgaa"
/label=DUS
2486. 2743
/feature="NMA0701"
/feature="NMA0701", possible pseudogene, len: 258 bp; shows weak similarity to part of SW:WBA_ECOLI (EMBL:M99166) wrbA, Escherichia coli Trp repressor binding protein (197 aa), fasta scores: E(): 0.064, 29.3% identity in 82 aa overlap. Also similar to part of TR:085450 (EMBL:AF067083) Vitreoscilla sp. Trp repressor binding protein (fragment) (124 aa), fasta scores: E(): 6.6e-14, 56.6% identity in 83 aa overlap
/codon_start=1
/pseudo
/transl_table=11
/product="hypothetical protein NMA0701 (pseudogene)"
complement(2755..2764)
/feature="Core DNA uptake sequence: gccgcgtctgaa"
/label=DUS
2778. 2787
/feature="Core DNA uptake sequence: gccgcgtctgaa"
/label=DUS
2861. 2864
/feature="Core DNA uptake sequence: gccgcgtctgaa"
/label=DUS
2871. 3530
/feature="NMA0702"
/feature="NMA0702", len: 219 aa; similar to TR:052925 (EMBL:Z50189), exsB, Rhizobium meliloti putative regulator of succinoglycan biosynthesis (not a transcriptional regulator) (234 aa), fasta scores: E(): 3.1e-15, 32.7% identity in 205 aa overlap. Also similar to many bacterial hypothetical proteins e.g. SW:IBA_HAETIN (EMBL:U032798), H11191, Haemophilus influenzae hypothetical

Query Match	Best Local Similarity	Matches 1282: Conservative	Score 92.6%	Length 311321	DB 1	Pred. No. 1.9e-294	Mismatches 62	Indels 0	Gaps 0
1	atgattaaatcaaaaaagcttaaacctgcacatcgccgagcgagacgagacgagtcggtt	1	92.6%	311321	DB 1	1.9e-294	62	0	0
55476	ATGATTAAATCAAAAAAGCTTAACCTGCACATCGCCGAGCGGACGAGCAAGCATCAT	55417							
61	taagaagcgccgagcattaccggaatcgcggtgtgttcgaggaataatgcccgtatgac	120							
55416	TATGACGGGCGCCGTCATTAACCAAGTCGCGTTCGTCGGAAGAAATATACCGTATG	55357							
121	ccctcgatgaagaatcaagaagcgatccgctcaaaaaagcgcaagtgctgtttgaagc	180							
55336	CCCTCGATGAAGATCAAGGACGGCATGCGCTCAAAAAAGCCCAAGTCTGTTTGAAGAC	55297							
181	aaaaagaatccggcgctgtgttactctgcgcgcgcttcagcgcaaaatcgccgcgattcac	240							

[illegible]

/organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 334 a 363 c 362 g 285 t
 ORIGIN

Query Match 90.2%; Score 1212.8; DB 6; Length 1344;
 Best Local Similarity 93.9%; Pred. No. 1e-286;
 Matches 1262; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

```

QY 1 atattataaataaagaagcttaaaccttcacatgcgagcggcgaagcggcgaagcggcgtt 60
Db 1 ATGATTTAAATCAAAAAGGCTTAATCTCCATCGCGGAGACCGGAGCAAGTCAATT 60
QY 61 taagagcggccggccatctacgaagtcgcgtcttgccgaagaatactgcggtatgcg 120
Db 61 TATGACGCGCGGCGCATTTACGGAAGTGGCGTGGCGGAGAAATGTCGGCATGCGC 120
QY 121 cccctgaatgaagaatgaagaagcgcgttcgaataaagaagccaaagtgcgttgaaagc 180
Db 121 CCGTCGATGAATAATCAAGGAGGAGCGCTCAAAAAGGCCAAGTGTGTTGAAGAC 180
QY 181 aaaaagaatccggcgtggtgcttactgcgcgcgcgttcgaacaaatccgcgcgtcac 240
Db 181 AAAAAGAAATCCGGCGGAGTATTACTGCGCGGCTTCAAGCAAAATCGCGCATATTAC 240
QY 241 cgtgagcgaagaagcgcgtactcagtcagtcggtgattgcgttgaaagcgaagcgaatc 300
Db 241 CGTGGCGAAAAGCGCGTACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 301 gaattgaacgtacgaacgtacgaacgtacgaacgtacgaacgtacgaacgtacgaacgt 360
Db 301 GAGTTGGAAGCGTACGTAAGTCTGAAGCGCTTGCAGAAATTTGACGACGAAAGTCCGCGC 360
QY 361 aaccctgaatccatccggttctggtgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 420
Db 361 AACCTGATTCATCAAGCGCTTATGACGCTTATGACGCTTATGACGCTTATGACGCTTAT 420
QY 421 ggcgtgattgcgaagcgcgttcgcacatctgcataatgcgaatgcgaatgcgaatgcga 480
Db 421 GCGGTAGATCCGAGCGGCTTGCCTATCTTGTCAATGCGATGAGACCAATCCGCTGCT 480
QY 481 ggcgaacccatcgcgtcattatcaaaagaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 540
Db 481 GCGGACCCCTAGCGTCAATCAAAAGGAGCGCGGAGACACTTCAAAACCGCGCTTGTGTA 540
QY 541 ttgagccgtttgacgaagcgaacatcgtttgtaagcagctggcgaagcgttcgcgcgcgc 600
Db 541 TTGAGCGCGCTGACCGAAGCTTAATTCATGCTGTGAAGACGACGACGACGACGACGACG 600
QY 601 tctgaagaatgctgcgaacatcgaacatcgaacatcgcgcgcgcgcgcgcgcgcgcgcgc 660
Db 601 TCTGAAGAATCTGCTCAATATCGAAACACATGAATTTGGCGCGCATCTCGCGGCTTG 660
QY 661 agtggcgcgcacatcattcatcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 720
Db 661 AGTGGCAGCAGCATTTCTTTTCATCGAGCGCGCGCGAGTAACCGATGATGATGATGATG 720
QY 721 aattatcaagaatgtaattacatcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 780
Db 721 AATTATCAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
QY 781 cgcgtgattgcgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttc 840
Db 781 CCGGTGCTTCCCTTGGCGCGCTGCAAGTCAACAAACCGCGCTTGTGCGATGATGATGATG 840
QY 841 ggtgcgaagaatcgaacatcgaacatcgaacatcgttcgttcgaacagaacacgcgcgtgalt 900
Db 841 GGTGCGAAGGCTTCTCAACTTACCGCGCGGAATTTGTTGACGCGCAACCGCGGTAT 900
QY 901 tccggttcggttatgaagcgcgcgttacacaaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 960
Db 901 TCCGCTTCCGTTATTGAACGGTTCGATGACAAAGCGCGCATGATTAATTTGGGACGCTAC 960

```

```

QY 961 cacaatcgaattccggtatcgaagaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1020
Db 961 CACATATGATTTCCGTTATTCGAAGAAGCGCGACCAAGAGCTTCTGCGCTGGCTTGGC 1020
QY 1021 ccgagccggagaacatcactcatcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1080
Db 1021 CCGGAGCGCGCAAAATCTCATCATCGCGCACCACTCTCGGCGCATTTCTTAAAAACAAA 1080
QY 1081 ctcttaagttcaacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1140
Db 1081 CTCTTCAAGTTTACGACAGCGCTTCAACGCGCGCGCGCATGATGATGATGATGATGATG 1140
QY 1141 taagagcgcgtgattgcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgttcgc 1200
Db 1141 TATGAGCGCGTAATGCGCTTGCATCATCTGCGCTTCTGCTTGTGCGCGATTAATGCTG 1200
QY 1201 ggcgtatccgacgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1260
Db 1201 GCGCATACCGACGCGCGCGCGCTTGGGTTGCTTGGATGAGACGAAGAACCTGCT 1260
QY 1261 ttgtagcgttcgcgttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1320
Db 1261 TTGTCAGCTTCTGCTTGCCTGCGCGCAAAATGCAATACGCGCGCTGTTGCGCAAGTGTG 1320
QY 1321 gaaacccattggaagaagcgtcga 1344
Db 1321 GAAACCATTTGAGAAAGGCTGA 1344

```

RESULT 7
 AE006171 10563 bp DNA linear BCT 08-MAR-2001
 LOCUS
 DEFINITION Pasteurella multocida PM70 section 138 of 204 of the complete genome.
 ACCESSION AE004439
 VERSION AE006171.1 GI:12721691
 KEYWORDS
 SOURCE Pasteurella multocida.
 ORGANISM Pasteurella multocida.
 Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 Pasteurella.

REFERENCE 1 (bases 1 to 10563)
 May, B.J., Zhang, Q., Li, L.L., Paustian, M.L., Whitlam, T.S. and Kapur, V.
 Complete genomic sequence of *Pasteurella multocida*, PM70
 proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)
 TITLE Direct Submission
 AUTHORS Zhang, Q. and Kapur, V.
 JOURNAL
 MEDLINE
 REFERENCE
 TITLES
 JOURNAL

FEATURES
 source
 1..10563
 /organism="Pasteurella multocida"
 /strain="PM70"
 /db_xref="taxon:747"
 94..1605
 /gene="PM1326"
 94..1605
 /gene="PM1326"
 /codon_start=1
 /transl_table=1
 /product="unknown"
 /protein_id="AAK03410.1"
 /db_xref="GI:12721692"
 /translation="MWIDVYTORDCILKLNKSPFDPVPLKNDVDIYAGEVHCY
 GENAGKSTLCKIAGIYPEOGEMFEGSKSYPKYKFAOSIGIGIHOELMVPOL
 SYLENFIEGKEKIVFGCKMWMQMRHKTQDQISETLIDPDKADVLSTADQMWVI
 AKAVFSRYKIITIDEPSSISKRNTQLEKIIHOLKKNVAMIIYSHRDEPOTIADO
 VTVLRDSIRGTMAVYRDTSPDEIVKIMWGEINFAEYHRYHLPTRVLRVESLTKHIT

REMARK The H. influenzae sequence has been updated by R. Fleischmann. New database matches have been assigned, product names have been improved, and a number of frame shifts have been corrected. We gratefully acknowledge the work of Tatusov et al. We have incorporated their annotation into the /notes fields of the corresponding H. influenzae genes

REFERENCE 5 (bases 1 to 13299)
White, O., Clayton, R.A., Kellavage, A.R., Fleischmann, R.D., Peterson, J., Hickey, E., Dodson, R. and Gwinn, M.
JOURNAL Direct Submission
TITLE Submitted (28-MAY-1998) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
REMARK The whole genome was shifted by 588 nucleotides for a new start On Sep 30, 1996 this sequence version replaced gi:1220812.
COMMENT Location/Qualifiers
FEATURES
source 1..13299
/organism="Haemophilus influenzae Rd"
/db_xref="taxon:71421"
complement(85..1455)
/gene="HI0161"
complement(85..1455)
/gene="HI0161"
/note="similar to SP:P06715 GB:M13141 PID:146248
PID:466637 GB:U00096 percent identity: 74.17; identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="glutathione reductase (gor)"
/protein_id="AAC21833.1"
/db_xref="GI:1573119"
/translation="MTKHVDYIAIGGSGGASLNRAAYGKCAIIEAKHIGTCVAVGCVPKKMYEYGAHIAEINNNYADYGFDEVKFEDESKLESROAYISRHTGEYNYLAKNNIDVINGFGFVADHTEVTLADGTEKOVADHLLTAGRPRNPNTKGGEGYIGDSGEFATLTELPRKAAVAGVAVGAVLSEVYNSIGVSHLLVRRHAPRNDPLIVFLLVYLAQDGIQHTNSPSELTNRKADSLTYRCDSGDYDPCYVAAAGRPETDKICLENAGVETNENGYKVDKDYKONTNKKYIYAGDITENKIELTPVAAAGRRISELFINRKPTETIDYSLAPVYFESHPRITGVGLTEPOATEQYGAENAVKVSSTFAMTAVTQHRKPTCKMLVCVGDKEKVLGHIGFVDEMIOGEFAVAILKMGATKADFDNTVAIRHTGSEEFVYMR"

gene
CDS
complement(1671..2270)
/gene="HI0162"
complement(1671..2270)
/gene="HI0162"
/note="similar to GB:U00096 PID:1773118 PID:1786637
percent identity: 25.82; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=1
/product="lipoprotein, putative"
/protein_id="AAC21834.1"
/db_xref="GI:1573120"
/translation="MALVNRKIKTLSSVGLIAUTLFLAGCAQNSIILAFPPAPASMN
VNRVAASVYTKDSRAIOELIASYKHGELIKLNSPSVTLFQOVMOONLISKFRVG
QLNSNMAWYDVREFTQVOEGNLRKLNKIQATYVVOGAKSYNSKFNVTSHOEG
VFNAGNDEIKHVLSQTFNDIYNNIYQOEVAALINQSN"
2364..2675
/gene="HI0163"
2364..2675
/gene="HI0163"
/note="similar to PID:663269 percent identity: 46.53;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="transcriptional regulator, putative (bola)"
/db_xref="GI:1573121"
/translation="MSIOQITEKTKQEFQPHFAIENESHLSNKGSESEFCVIV
SADPKNIRKQVRHORITQLNEELNHSIHALLHLPPEEKAKQNETVPHSTKAGIG
R"
2966..4309
/gene="HI0164"
2966..4309
/gene="HI0164"
/gene="HI0164"

gene
CDS
complement(1671..2270)
/gene="HI0162"
complement(1671..2270)
/gene="HI0162"
/note="similar to PID:663269 PID:663271 percent identity:
71.33; identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="NADH:ubiquinone oxidoreductase, subunit B
(ngB)"
/protein_id="AAC21837.1"
/db_xref="GI:1573123"
/translation="MGLNKLFEKMEPAFLPGKYSKLYPIFESYITLTYPTGYTHKN
THVBDALDSKRMKMTVFLALPFAIEGYWVNGAIPALNOLGNDQIADWNYALA
SSLDLITANNTWGSKMGALGIFLPIYLVVFTVCTITELLESYVRGEHVEGMVST
ILFLIYVPTPLPMDOALGIFGIIYVAIEIFGVGRNMPNLAGRALFPAYPAOIS
GDVPTADGSGGADALSOWSGGALQHVTVGAPITWMOAFNGNLGSGEYGTALA
ILIGAVYFRIAMRILAGVMIGMTATSTLFINLSETPMFSPMHMHFVYGA
LGMVMAADPVSASTNGKMYTGALLGVMAVLIRTVPAIPEGMALILRANLFAPI
FDIYVQANIKRRKARTNG"
5540..6274
/gene="HI0167"
5540..6274
/gene="HI0167"
/note="similar to PID:663269 PID:663272 PID:893413 percent
identity: 54.36; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=1
/product="NADH:ubiquinone oxidoreductase, Na
translocating"
/protein_id="AAC21838.1"
/db_xref="GI:1573124"
/translation="MAKPNKDSVGGTIIIVLLISVCSIIYAGSNVKKPNOBOKIL
DKOKILVAGLQENNTVKEFYAKFIEPRFVTLATGETTOADDSQOATPRADAKAR
IRSRKRTTEVYLVADGQOTQVILPIGTGLMSVYGLVSVQPDGNTINGITYOHG
ETPLGGEIENPMNNAISLEKGRKLPDEQHOPAIRYKGOAPODEHSIDLSGATLNG
VOGTENYFSSDGFPGYLEKLSGAN"
6274..6900
/gene="HI0168"
6274..6900
/gene="HI0168"
/note="similar to PID:663269 PID:663273 PID:893414 percent
identity: 74.88; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=1
/product="NADH:ubiquinone oxidoreductase, Na
translocating"
/protein_id="AAC21839.1"
/db_xref="GI:1573125"
/translation="MSGKTSYKDLLADIAKNPILQIILGICALAVATYTKLEAFVM
AIAVLEVYLSNLEFVSLTRNYIPNSIRIYOLAIYASIVYVDOLIKAVYAGLSKOLS
VFVGLIITNCIYMGRAEAFAMKSPVESFGLGGLGYGMLIIVAFRELLISGKIL
FGMTIFERTIONGMYOANGFLPLAPSAFIIIGFVIWGLRTWKPEOQEK"
6904..7500
/gene="HI0170"

[illegible]

AUTHORS
Stover, C.K., Pham, X.O., Erwin, A.L., Mizoguchi, S.D., Warren, P., Hickey, M.J., Brinkman, F.S., Hutnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltzy, L., Tolentino, E., Westbrook-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., and Paulsen, I.T.
Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen
Nature 406 (6799), 959-964 (2000)

JOURNAL
MEDLINE
20437337
2 (bases 1 to 16903)

REFERENCE
Stover, C.K., Pham, X.O., Erwin, A.L., Mizoguchi, S.D., Warren, P., Hickey, M.J., Brinkman, F.S.L., Hutnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltzy, L., Tolentino, E., Westbrook-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., Paulsen, I.T., Reizer, J., Sater, M.H., Hancock, R.E.W., Lory, S. and Olson, M.V.
Direct Submission
Submitted (16-MAY-2000) Department of Medicine and Genetics,
University of Washington Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA

TITLE
JOURNAL
Location/Qualifiers
1. 16903
/organism="Pseudomonas aeruginosa"
/strain="PAO1"
/db_xref="taxon:287"
complement(67, .2292)
/gene="PA2984"
complement(67, .2292)
/gene="PA2984"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAG06372.1"
/db_xref="GI:9949084"

FEATURES
source
1. 16903
/organism="Pseudomonas aeruginosa"
/strain="PAO1"
/db_xref="taxon:287"
complement(67, .2292)
/gene="PA2984"
complement(67, .2292)
/gene="PA2984"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAG06372.1"
/db_xref="GI:9949084"

gene
2423. .2956
/gene="PA2985"
2423. .2956
/gene="PA2985"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAG06373.1"
/db_xref="GI:9949085"

CDS
/translation="MPRLKRYMPDESIRSOQLRGLKLTHTDPMIHLNRHSVR
AMAAGLFAAIPMPOMLILAGLAVIRSNLPISSGLVMTNPMPVPCYKILA
WLAQVPRRLPDHLTWEMISQELSLMQPPLGASVCGIATLITMLYMRWVO
RSMRRCRORRERVAAK"
complement(304, .4345)
/gene="PA2986"
complement(304, .4345)
/gene="PA2986"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAG06374.1"
/db_xref="GI:9949086"

gene
2423. .2956
/gene="PA2985"
2423. .2956
/gene="PA2985"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAG06373.1"
/db_xref="GI:9949085"

CDS
/translation="MPRLKRYMPDESIRSOQLRGLKLTHTDPMIHLNRHSVR
AMAAGLFAAIPMPOMLILAGLAVIRSNLPISSGLVMTNPMPVPCYKILA
WLAQVPRRLPDHLTWEMISQELSLMQPPLGASVCGIATLITMLYMRWVO
RSMRRCRORRERVAAK"
complement(304, .4345)
/gene="PA2986"
complement(304, .4345)
/gene="PA2986"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAG06374.1"
/db_xref="GI:9949086"

gene
2423. .2956
/gene="PA2985"
2423. .2956
/gene="PA2985"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAG06373.1"
/db_xref="GI:9949085"

CDS
/translation="MPRLKRYMPDESIRSOQLRGLKLTHTDPMIHLNRHSVR
AMAAGLFAAIPMPOMLILAGLAVIRSNLPISSGLVMTNPMPVPCYKILA
WLAQVPRRLPDHLTWEMISQELSLMQPPLGASVCGIATLITMLYMRWVO
RSMRRCRORRERVAAK"
complement(304, .4345)
/gene="PA2986"
complement(304, .4345)
/gene="PA2986"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAG06374.1"
/db_xref="GI:9949086"

ALRRREVGAAPFAELQGMUSYKGMPLVYNGIDPOEERKVSIVGEHIVOGSLDDIK
PGEHIVGEITARRFHVWGDJLTLVPEVTSAPGITPMOHTTVALEKVAELD
NSLIDIDAGOLRLPOGVSVLELIDYOSPOVAKVAVLELOGFSSTWPT
OGSEFNKMEKEMIGILLIILVAVFENIATILINVADKRTDIALRTGAPPROI
MAIMVOGTIVIGTIVIGVAGFALNITGMIDRIERLVGHVFSDDVFIYVLS
DLQYLDVILCSAALMSPLATLTPSRRAARTQAEALRE"
complement(4358, .5041)
/gene="PA2987"
complement(4358, .5041)
/gene="PA2987"
/codon_start=1
/transl_table=1
/product="probable ATP-binding component of ABC
transporter"
/protein_id="AAG06375.1"
/db_xref="GI:9949087"
/translation="MNDKSVLSCKNSLSYDEGPOVYVSGVELNLPGERVAIVS
SGSKSTLLMGLDTPSAGSVYLACEELSALEKVAHSHVAVAFPTOMQ
FTALENCPMLLIGRTPIAERORAELELRYGSHRSKPAELSGGEROVAIARA
LYNRPOLVILDEPGRNDQHTAGQIOELMELSLQTSFLVYTHDQLAGHMDRIAR
LEGRITAA"
complement(5034, .6284)
/gene="PA2988"
complement(5034, .6284)
/gene="PA2988"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAG06376.1"
/db_xref="GI:9949088"
/translation="MFRPLSVFTIGRTYTRAKRSHFVSLTSMIGALGVMTIV
LSVANGFDREMTIRILGMVQAVESYOPIDDMRLAEKVAHSHVAVAFPTOMQ
LSANGVQTVAVDADPLEAKVSIIDPEFGGSLAEKGGFIVIGOLAOKLGG
IGDKVTLIPVAVTPAGVPEPRMKRFIVYVFGAELGGSLIHLEDAARLOMK
TNOYQRIKIKLDLFOAPRVAMETARTLNDPVARMTSHGLYQAIRMEKICIL
HLITVAANFNITSLVAVYTDKSDIATILRTIGATPCQMATFMVQGYTGYTIGL
VGLVGVAAALNVSAMISALEKILGHQFLASDVFTIDLPQMLDVMVYVCGAIVL
SEFATILVPAARMAARTQPAELRYE"
6222. .6986
/gene="PA2989"
6222. .6986
/gene="PA2989"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAG06377.1"
/db_xref="GI:9949089"
/translation="MAATFGRIARADENGSGSERVALSMOKMDASGMRSGEASPAL
HSDHCRGAPMSTRDADREYVRIEDPTIALFERPAGVADGILESPRGATLPSLSD
LHIMDYSOHLRIHSRDRILANYIKVMKRRIDLLQVYVQSLKEIGERKYSLSF
GGVSEFRDRAIPVQQLVLRNVLPGFGELRARKVIAHOHDEFETIGTEFALSDA
QROLARHIILOKQAOORRLAAGGPGEGPQPSST"
7318. .8040
/gene="PA2990"
7318. .8040
/gene="PA2990"
/codon_start=1
/transl_table=1
/product="probable phosphodiesterase"
/protein_id="AAG06378.1"
/db_xref="GI:9949090"
/translation="MTLYGHRGAKGAPENTLASFORCLEHGVNRCELDILSRDGE
LMVTHPTKRTTGRGKVAHDAATATADARAGGVAWPPPIPLNELFEKCPPE
HMOLEYSASRERAAPTVYTRLETTERPGNDKYTVSGSTVYKIPANLELPSRL
VAETAMDPLKVAOHYCALALAMWTICTERILKAKOROGILHVSVMYVNEPALMRILA
DFGVDSLITPFGLASATLGNR"
complement(8083, .9477)
/gene="sth"
/note="PA2991"
complement(8083, .9477)
/gene="sth"
/codon_start=1
/transl_table=1

Db 1048 CGTGGGCTAAAGTGTCTCAATCTGTAGTGAATGAAGGCGAAGGAGAGAGTG 1107
 Oy 301 gaatttgaacgcaacgcaaccgtgaagcgctgcaactaaagcggaagagtgccgcgc 360
 Db 1108 ACATTGATAGTTCGAAGCGGCTCAACTTTTCAGTCTAGTGTGAAGTATCAAGACT 1167
 Oy 361 aacctgcaatccggttctgtgagctgcctgcgacccgctcgcgttcagcaaatctcc 420
 Db 1168 CAGTGTGACTCTGCGCTATGAGCCGCTTTAGCTACTCTCGTTCGTTAGCAAGTTCT 1227
 Oy 421 gcccgtgagtcgagcgccgttcgcacattcgtcaatgcagggcaccacccgtgctc 480
 Db 1228 GCAATGAGTCTTCTACTAAGGCTATTTCCTAATGCAATGATACATTAATCATTAGCA 1287
 Oy 481 gcccagccctacggttcattatcaaaagcgcgcgagatctcaaaagcgccgttctgta 540
 Db 1288 GCTAAGCGCTGATGATCAATTAAGCAACAGACATTCATGCTGTCTAGACATT 1347
 Oy 541 ttgaagccgttgacgcaagcaaaatccatgcttctgtaagcgagctggcgacgctgcg 600
 Db 1348 CTTTCAGCCCTGACAGAGGCAAGGTTTACGTATGTA--ATCTGGCACTAGTTGCCA 1404
 Oy 601 tctgaaatgctgcacacatcgaaacatgaatctcgcgccgcgcctcctgcgcttg 660
 Db 1405 CGCTCTTACAGTCTAAGCTAAGAGAACAGTCTTCGATGGCCCTCACCCAGCAGTCTT 1464
 Oy 661 agtggcagcaacattcatctacgcgctgcgcgcgagataaaacgctgtgacacalc 720
 Db 1465 GCTGGCAGCCCATGATGCTTCTTACCACTGATGATGAGAAAGTGGCAAGGAGCAAT 1524
 Oy 721 aattatcaagatgaattacatctgcccgttctgttgcacagcgccgctgcagaccgag 780
 Db 1525 AACTACCAAGAGCTTATCGCTTGCTGTAAGCTGTCTTACTGCTGAGCTTATACACAGAT 1584
 Oy 781 cgcgtgattgcccagtgagtgctcgaagtcacaaacgcgcctctctgctacccgtt 840
 Db 1585 CGTGTGTGCTCTAAGCTGCTGCAAGTGTGATTAACCCGCTGATGTTGCCGTAATC 1644
 Oy 841 ggtgcgaagatgcgcaaatctacgcgagcgatgtgtgacacagacaaacgcgctgatt 900
 Db 1645 GGTGCAAGCTTGATGATCACTTAACAGATTAACAGTGTGATGCCCTGGGCAAGTTGCTGAT 1704
 Oy 901 tccggttcgagatgaagcgcgcggttacaacagcgcgacagatatttggagcctac 960
 Db 1705 TCTGCTTGTGATGACAGGTAACATGCACTGCTCCACAGCATACCTTGCCGCTTAC 1764
 Oy 961 cacaatcagattccgttatcgaagaagcgcgcaagcaagaagcgtctcgtctgtgtgcg 1020
 Db 1765 CATCAACAGGTTTCTGATTAACGTGAGGCCCTGAGAAAGAACTATTGGTGGCAATG 1824
 Oy 1021 ccgagcgcgcaaatctacatccgcgttacaaacccctgcgcattctctgaaaaacaaa 1080
 Db 1825 CCTGTAAAGATTAAGTTCTGCTTAACCTGCTTCTTCCCTTGGTACACTATTCAAGGCTCAG 1884
 Oy 1081 ctcttcaagatcaaacagcgctgaacgcgcgacccgcgcacatgtgtgcgattgttact 1140
 Db 1885 TTGTTCAATATGACAAACAGCAACCAATGATGATGATGATGATGATGATGATGATGATGAT 1944
 Oy 1141 taagagcgctgtgattcccttggatattccctgcaacccctgcttcttgcgatttaactgc 1200
 Db 1945 TACGAACCGGTAATGCTCTAGATATGAGCCTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2004
 Oy 1201 ggcgataccgacagcgcgagcagcattggtgtctgtgaaattgagcaagaagcctcgc 1260
 Db 2005 GGTGTATACAGATGCTCTCAACACTGGGTGGCTTGAATGAAGTATTAAGCA 2064
 Oy 1261 ttgtgcagcttgcctgcgcgcaaatcaatacgcgacccgctgtgtgcgcaagctgtc 1320
 Db 2065 TTGTGTACTTGTATGTCTCAGGTAAAGTACGAGTACGATACGATACGATACGATACGATACGATAC 2124
 Oy 1321 gaacacattgaaagaaagcctga 1344

Db 2125 GATACTATGAGAGAGAGGTA 2148

RESULT 15
 A96090 474 bp DNA linear PAT 07-SEP-2000
 LOCUS Sequence 123 from Patent WO924578.
 DEFINITION
 A96090
 VERSION A96090.1 GI:6779937
 KEYWORDS
 SOURCE
 ORGANISM
 unclassified.
 unclassified.
 unclassified.
 1 (bases 1 to 474)
 AUTHORS
 TITLE
 JOURNAL
 PIZZA M., Scarlato, V., Rappuoli, R., Grandi, G. and Mesigian, V.
 Neisserial antigens
 Patent: WO 924578-A 123 20-MAY-1999;
 PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);
 CHIRON SPA (IT); GRANDI GUIDO (IT); MESIGIANI VEGA (IT)
 FEATURES
 source
 1..474
 /organism="unclassified"
 /db_xref="taxon:32644"
 BASE COUNT 122 a 130 c 130 g 91 t 1 others
 ORIGIN

Query Match 35.2% Score 473.6; DB 6; Length 474;
 Best Local Similarity 99.88; Pred. No. 1.8e-105;
 Matches 473; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 atgataaatacaaaaaaggtctaaactgcccacgcggygcagacggagcaagcgcgt 60
 Db 1 ATGATTAAATCAAAAAAGGCTTAACCTGCCATCGCGGCGACACCGAGCAACCCGTT 60
 Oy 61 taacgagcgccgcgcattaccgaagtgcggttctgttgcggaagataatgcggtatgc 120
 Db 61 TACGACGCGCCGCGCATYACCGAAGTCCGTTGCGGAAAGATATGCGGTATGCGC 120
 Oy 121 cccctgataaagtcaaggaagcgatgcgtcaaaaaagccaaagtctgttgaagac 180
 Db 121 CCTCGATGAAGTCAAGAGAGGCGATGCCCTCAAAAAAGCAAGTCTGTTGAAGAC 180
 Oy 181 aaaaagaatccggygcgtgtgttactgcgcgcgtcgaagcaaatgcgcgattac 240
 Db 181 AAAAAGATCCGGCGGTGTTTACTGCGCGCTTCAGCAAAATGCGCCGATTCAC 240
 Oy 241 cgtgcgaaaaagcgctacttcagtcaagtcgtgtatctgcgttgaagcgaaacgaatc 300
 Db 241 CGTGGCGAAAGCGCGTACTTACGTACGTGATTTGCGTTGAAGCAACGAAATC 300
 Oy 301 gaatttgaagctacgcacccgaagcgctggaacttaagcgcggaagatgcgcgc 360
 Db 301 GAGTTTGAACCTTACGCACTGTAAGCGCTGCAAACTTAAGCGGGAAGAAATGCGCGC 360
 Oy 361 aacctgataatccggttctgtgacgtgcgtgcacccgctcgttcagcaaatctc 420
 Db 361 AACTGATCAATCCGCGGTTTGAGACTGCGGCGCACCCGCTTCAGCAAAATTCCT 420
 Oy 421 gcccgtgagtcgagcgcttgcacattcttcgcaatgcatgtgaacacacacgc 474
 Db 421 GCCGTGATGCGGAGCGCTGTGCGCATCTTCGTAATGCGATGACCAACCAATCCG 474

RESULT 16
 AB008030 6908 bp DNA circular BCT 15-JUN-1999
 LOCUS
 DEFINITION
 Vibrio alginolyticus genes for Na-translocating NADH-quinone
 reductase complex, nqr operon, complete genome.
 ACCESSION
 AB008030 D43958 D49364
 VERSION
 AB008030.1 GI:2558472
 keywords
 Nqr1 subunit of Na-translocating NADH-quinone reductase complex
 alpha-subunit; Nqr2 subunit of Na-translocating NADH-quinone

SOURCE
ORGANISM
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.
1 (sites)
Hayashi, M., Hirai, K. and Unemoto, T.
Cloning of the Na(+)-translocating NADH:quinone reductase gene from the marine bacterium *Vibrio alginolyticus* and the expression of the beta-subunit in *Escherichia coli*
FEBS Lett. 356 (2-3), 330-332 (1994)
95104444

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
2 (sites)
Beattie, P., Tan, K., Bourne, R.M., Leach, D., Rich, P.R. and Ward, F.B.
Cloning and sequencing of four structural genes for the Na(+)-translocating NADH-ubiquinone oxidoreductase of *Vibrio alginolyticus*
FEBS Lett. 356 (2-3), 333-338 (1994)
95104445

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
3 (sites)
Hayashi, M., Hirai, K. and Unemoto, T.
Sequencing and the alignment of structural genes in the *nqr* operon encoding the Na(+)-translocating NADH:quinone reductase from *Vibrio alginolyticus*
FEBS Lett. 363 (1-2), 75-77 (1995)
95246889

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
4 (sites)
Hayashi, M., Nakayama, Y. and Unemoto, T.
Existence of Na+-translocating NADH:quinone reductase in *Haemophilus influenzae*
FEBS Lett. 381 (3), 174-176 (1996)
96176316

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
5 (sites)
Nakayama, Y., Hayashi, M. and Unemoto, T.
Identification of six subunits constituting Na+-translocating NADH:quinone reductase from the marine *Vibrio alginolyticus*
FEBS Lett. 422 (2), 240-242 (1998)
98149659

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
6 (bases 1 to 6908)
Hayashi, M., Unemoto, T. and Sugiyama, A.
Direct Submission
Submitted (09-OCT-1997) Maki Hayashi, Faculty of Pharmaceutical Sciences, Chiba University, Laboratory of Membrane Biochemistry; 1-33 Yayoi-cho, Inage-ku, Chiba, Chiba 263, Japan
(E-mail: makiha@p.chiba-u.ac.jp, Tel:043-290-2932, Fax:043-290-3021)
On Oct 22, 1997 this sequence version replaced gi:677955 gi:893412.
D43958: submitted (17-Dec-1994)
D49364: submitted (17-Feb-1995).
Location/Qualifiers
1..6908
/organism="Vibrio alginolyticus"
/db_xref="taxon:663"
/note="There are conflicts between this sequence and Genbank Accession Number:J237111."
386..1726
/gene="nqr1"
386..1726
/gene="nqr1"
/note="contains FMN as a cofactor; nqrA, Nqr1 subunit of Na-NQR complex alpha-subunit"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="Nqr1 subunit of Na-translocating NADH:quinone reductase complex alpha-subunit"
/protein_id="BAA22910.1"
/db_xref="GI:2558473"
/translation="MITTKGGIDLPDIAGTPSOVINDKTIKRVALLGEEYVGMPTMHVRDEVKKAQVLEEDKRNPGYKFTAPAGKVIENRGAKRVLOSIVLIVAGEQCVTFDKFPAQISGIDREVIKTQLVDSGLMTALRTRFPSKVPRIESTKAIFVTADTNFLA

COMMENT
reductase complex; Nqr3 subunit of Na-translocating NADH:quinone reductase complex gamma-subunit; Nqr4 subunit of Na-translocating NADH:quinone reductase complex; Nqr5 subunit of Na-translocating NADH:quinone reductase complex; Nqr6 subunit of Na-translocating NADH:quinone reductase complex beta-subunit.
Vibrio alginolyticus DNA.

FEATURES
Source

gene
CDS

conflict
/gene="nqr1"
/citation=[2]
/replace="c"
1234
/gene="nqr1"
/citation=[2]
/replace="t"
1732..2976
/gene="nqr2"
1732..2976
/gene="nqr2"
/note="hydrophobic membrane protein with pi 8.14; nqrB, Nqr2 subunit of Na-NQR complex"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="Nqr2 subunit of Na-translocating NADH:quinone reductase complex"
/protein_id="BAA22911.1"
/db_xref="GI:2558474"
/translation="MALKKFLIEDIEHFEPPGKHEKFEALYEVATVPTPGVTNKS SHVRSDVLKRIIMVWFVAVPMPGMYNAGGQAIALNIHMYAGDOLATVLSGMHY WLTEMLGGTIAADAGVSGKMLGATVPIVATVPLVNGSFVPLPEVMYKHEVNEFF VTSITFALIVPEPLPLNQALGITTGVYAKITFGTGRNPLPLAGAFLEFPAI QISGDVYVTAADGSGATALSQAGNGALVNTVTSPTWMDATIGNIPSGISGV TILADIGAMATVYMRILASWRIIAGVIMIAVSTLNVIGSDTNPFNPMWMLVILG GFARFEMFMAIDVPVASFSTNKGKMYGILIGAMCYMIRVNPAYPEGMALILFANLF APLFDHVIVIERIKRRLARYCK"
181
/gene="nqr2"
/citation=[2]
/replace="c"
2400
/gene="nqr2"
/citation=[2]
/replace="c"
2417..2418
/gene="nqr2"
/citation=[2]
/replace="c"
2614
/gene="nqr2"
/citation=[2]
/replace="g"
2966..3736
/gene="nqr3"
2966..3736
/gene="nqr3"
/note="acidic protein with pi 4.83; nqrC, Nqr3 subunit of Na-NQR complex gamma-subunit"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="Nqr3 subunit of Na-translocating NADH:quinone reductase complex gamma-subunit"
/protein_id="BAA22912.1"
/db_xref="GI:2538475"
/translation="MASNDSIKRTLGIVIGLSVCSIIYSTAAVGLRDKOKANAVLD KQKRIEYVAGIDANGKRVPELFAEYIEPRVLDLETGNETEGNASTYDQREASDKERS IALTPEEDVADIRRRANTAVVYLKDDDEVOKVILPMHGKLSMMAFVAVETDGNF VSAITLYEGGETPELGGEVENPSRDOFIGIKLYNEDHOPAIKYVKGGAPOGSEHGVD GISGATILTSNGVQHTFPFWLGDGKGFPLAKVRDELM"
3295
conflict

Victoria 3052, Australia
 Location/Qualifiers
 source
 1. 1356
 /organism="Porphyromonas gingivalis"
 /strain="W50"
 /db_xref="taxon:837"
 1. 1356
 /codon_start=1
 /transl_table=11
 /product="50 kD antigen Pg1"
 /protein_id="AAD33930.1"
 /db_xref="GI:4929280"
 /translation="MAKVKTKKGLALNKGKPLPMLAEPAPSYAVYVDDPEGYI
 KVTARPDGKVRAGSALMHKAYPEKFTSPVSEGYIANKKRYLSIEVPDGLN
 EESPVPDPSALSAEOLIKELLISGMWFIKQRPDIATPIADPIDITANFTAP
 LAPDFEIVRGERALQFAIDALAKLTGKVVGLPGSSLSLHMEIVHGHHPAG
 NMGVLNHTKRIINRGFTVMTLKNATDLIVGRFLTGKADFTRMILAMTGSDAAHGYR
 IMSGNVFASPRGRITKESHERVIDGNVLTKGLKEKEPELSARCDQIVPEGDV
 DELFGMAAPRLQYSMSRAYFSLGSKNKEYVDARIKGERAMINSEYDRVPMNI
 IPEYLKALIAFDIDMEDGLIEVAPEDPATCEFYDTSKIEIQRIVREGIDMLYKEM
 N"

BASE COUNT 338 a 333 c 380 g 305 t
 ORIGIN

Query Match 8.5%; Score 113.6; DB 1; Length 1356;
 Best Local Similarity 46.1%; Pred. No. 3.6e-17;
 Matches 623; Conservative 0; Mismatches 699; Indels 30; Gaps 6;

QY 2 tgattaaatcaaaaaggtctaactgcgcacatcgcgagcgagcgagcaagcgttt 61
 Db 11 TTTATAAAACAAAAGGCTTGCACCTTAATCGAAGGAAAACCGTGCCGAGATGC 70
 QY 62 acgacgccccgccc--attaccgaagtcgctgtcttgcgcaagaatatacgcgatgc 118
 Db 71 TGGCCGACCGCCCAAGAGCTCTACTACCGCGTGGTCCGACGATTTGAAGGTGTTA 130
 QY 119 gccccctcgatgaagaagtcgaagcgatgcgcgtcaaaaaagcgcaagtcgtttgaag 178
 Db 131 TCCCAAGGTCGACGGCTGTCGGGGGATTAAGTGCGTGCCGCTCACACACTGATGCACC 190
 QY 179 acaaaaagaatccgagcgctgtcttactgcgcgcgcttcacagcaaaatccgcgagatlc 238
 Db 191 ACAAGGATATTCGGAGATATAAGTTTACAAGTCCGGTTAGCGGGAAGTGAATCGCGGTGA 250
 QY 239 accgttgcgaaaagcgctactcattcagtcagtcgtgattgcgcttggaagcaagcagaa 298
 Db 251 ATCGCGGTGCCAAGCGCAAGGTTTGAGCATCGAGGTGMAACCGGACGACTGAAGCAAT 310
 QY 299 tcgagtttgaacgtctacgacactgaagcgctgcgaacttaagcgcgagaagatgcgcc 358
 Db 311 ACAGAT-----CATTCCTCTGTCGGGATCCGTCCTCTCTCCGCAACGATCAACG 364
 QY 359 gcaacctgataccatccggtttgtagactgcctgcgcacccgctcgccttcagcaaaatlc 418
 Db 365 AGCTTTTACGTGCAAGGGATGTGTGGGGTTTATTAAGCAACGTCCTTAACGACATAGTGG 424
 QY 419 ctgcgctcgatgcgagcgcttcgcacatcttcgcaatgcatggaacccaatccgcgtg 478
 Db 425 CTACACCGGATATGATGCTCCACGCGCATTTATTTACTGCCAACTTTATGCAACCATTTGG 484
 QY 479 ctgcgacccccaagcgtatatacaagaagcgcgaggaattcaaaacgcgctgtttg 538
 Db 485 CTCGGGACTTGATTTTATCTGTCGAGAGAAAGACGCCCTTGCAAGACTGCAATGATGATG 544
 QY 539 tattgaacgcttgacgcaagcaaaatccatglttgaagcagctgcgcgacagctgc 598
 Db 545 CCTTGGCCAAACTCAGCAGAGAAAGGTGTATGTGGGCTTGAAAGCCGGGCTTCACTCTCG 604
 QY 599 cgtctgaanaatgcgtccaaatcgaaacatgaattcggcgagccgcgactcctgcgcgtt 658
 Db 605 GCTTGCAATGACGAATGCTGAGATGACA-----CGGACCTCATCCGGCAGGTA 655

QY 659 tgagtgcacgcacattcatttcacgagccgctgcgcgcgcaataaaccgltgtggacca 718
 Db 656 AGCTGGGCGTGTGATCAATCATACGACACCCCAATCAATGAGGGCGAAGCGTGTGACGC 715
 QY 719 tcaattcaagaatgtaattaccattgcccgtttgttcaacagcgctcgttaaacacg 778
 Db 716 TCAAGGCTACCGACCTGATCGATCGATCGAGCTTCCGTGTACGGGCAAGCCGATTTTA 775
 QY 779 agcgcgtatgctccctagtggtttcctcaagtaacaacacgcgcctctgctgacggtt 838
 Db 776 CCAGATGATGTCATGACCGGCTCGAGCGTGCACCTCACGATVAGTCCGATTTATGTC 835
 QY 839 tgggtgcgaagatcgcgaattacgcgcgcgcgaattg-----gtttacacagacaacc 892
 Db 836 CGGGTTGCAATGTCTTTCCTTCCCGCGGACTGACATTAAGGATTCACAGC 895
 QY 893 gctgattccgcttgatgattgaagcgcgcaatcaacaagcgcgacagattatttg 952
 Db 896 GTGTGATCGATGCAATGTGCTGACCGGTAAAGACTCTGAGAAAGACCTTTCCTGT 955
 QY 953 gaagctacacaatcagattccgctatcgaaagcgccg----agcaagagctgtgtc 1009
 Db 956 CAGCCCGTGTGACACGATCACGCTGATCCCGAAGCGACGATGTGACGACTTTCG 1015
 QY 1010 gcttggttgccgcgcgacgcgaacaatactcaacgcgtacacacctgycacattcc 1069
 Db 1016 GGTGGCTGACCCCTCTTCGATCAGTACGATGACGATGACGATGATTTCTGTGTTGC 1075
 QY 1070 ----tgaaaaaacatcttcgaattcaacaacagcgctcaacgcgcgcgcgcgcgtatg 1126
 Db 1076 AGGGGAAAAAAGAGTACGTACTGATGCCGAGTCAAGGGTGGGAACGTGCTATGA 1135
 QY 1127 tgcgatttgtaacttcgcgcgcgctgtagtccttgatataccctgacacctgtcttgc 1186
 Db 1136 TCATGACACAAGATATGACCGCGCTTTTCCGATGACATCTATCCGAGATTTGCTCA 1195
 QY 1187 gcgatttaactcgtcgcgcgataccagacgcgcgcgaagcatttgggttcttgaaatlgacg 1246
 Db 1196 AGGCTATTTATGACATTCGACATGACGACGAGTGAAGATGAGCATATATGAGTGGCTC 1255
 QY 1247 aagaagacctgccttgcgtgaagcttcgtgcgcgcgcgcgcgcgcgcgcgcgcgtgt 1306
 Db 1256 CGAGAGACTTTGGCACTTGGCAATTTGTGATATCAATCAAGATGACGCTGACGCTATCG 1315
 QY 1307 tgcgcaagtcgtggaacatttgagaagaa 1338
 Db 1316 TTGCGAGGGCTTGATATGCTCTATAAGGA 1347

RESULT 19
 HITU20229 4926 bp DNA linear BCT 09-FEB-1995
 LOCUS HITU20229/c
 DEFINITION Haemophilus influenzae BOLA (bola), glutathione reductase (gor),
 phosphatidylserine decarboxylase (psd), 30K protein (rpfm), genes,
 complete cds.
 ACCESSION U20229
 VERSION U20229.1 GI:644850
 KEYWORDS
 SOURCE Haemophilus influenzae.
 ORGANISM Haemophilus influenzae
 Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 Haemophilus.
 1 (bases 1 to 4926)
 AUTHORS Barcak,G.J. and Heimer,S.R.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 4926)
 AUTHORS Barcak,G.J.
 TITLE Direct Submission
 JOURNAL Submitted (24-JAN-1995) Gerard J. Barcak, Biochemistry, University
 of Maryland at Baltimore, 108 North Greene Street, Baltimore, MD
 21201, USA
 FEATURES
 source
 1. 4926
 Location/Qualifiers

organism="Haemophilus influenzae"
/strain="Rd"
/isolate="Kw20"
/db_xref="taxon:727"
complement(27..224)
/note="orf65"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AA62134.1"
/db_xref="GI:644851"
/translation="MITIKKDLPIAGKPAOVHSGNAVNOAILGEEYGMPSMK
VREGDYVKKGOVLEDEKILY"
complement(519..842)
/gene="bola"
complement(519..830)
/gene="bola"
/function="stationary phase regulatory protein"
/codon_start=1
/transl_table=11
/product="bola"
/protein_id="AA62135.1"
/db_xref="GI:644852"
/translation="MSIQIIEOKIOKEFOPHPLAIENSHLHNSNGSSEHPCVY
SADFKNIRKYORHRIYQILNEIINSHIALALHLFTPEEMKQNETVPHSTKACGIG
R"
misc_feature
775..785
/note="encodes DNA uptake recognition sequence"
complement(838..842)
/gene="bola"
877..913
/note="OxyR recognition sequence (Tartaglia, L.A., et al.,
J. Mol Biol. 210: 709-710, 1989)"
/bound_moiety="OxyR"
915..919
/note="orf199"
924..1323
/note="orf199"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AA62136.1"
/db_xref="GI:644853"
/translation="MALVVKITLSSVCIILATILFLAGCAQSNITLFTPPASAKN
VNRPAVSVTKDSRAIQETASTYKHELLKLNASPSVTQLFOQMOONLSKFERG
QUNGSAWVTVDFEFGTQVQGNLRKLTKTQATVYVGAGSYNKSFNVHSQEG
VFNAQNDIHKVLSQTFNDIVNNIYQDOEYVAAAINOYSN"
1343..3109
/gene="gor"
1343..1348
/gene="gor"
/note="putative"
1365..1370
/gene="gor"
/note="putative"
1585..1590
/gene="gor"
/note="putative"
1605..1610
/gene="gor"
/note="putative"
1728..1731
/gene="gor"
1739..3109
/gene="gor"
/EC_number="1.6.4.2"
/function="reduces oxidized glutathione using NADPH"
/codon_start=1
/transl_table=11
/evidence="experimental"
/product="glutathione reductase"
/protein_id="AA62137.1"
/db_xref="GI:644854"

/translation="MKHYDYLAIIGSGSGIASINRASTGKCAITIEAKHLGCTCVN
VGCVPKVMFEGAHIAEINNYPADYDFDEVKKFDPSKLESQATISHIHYSNV
LAKNNIDVINGFEKFDVAHTLEVTIADGTEQVYADHILLATGSRPYRPNIKQOEXGI
DSDEFAITELPKRAAVIGAGYIAVELSGYNSIGVTHLYVRHARMRODPLIVET
LVEIYADQGIOLHNTKSTPSEIVKAKGSDIVRCGSGSDVDCVYAAAGRPPTDKIG
LENAGVETNEHYKVDKQNTNKGIGYAAVDIENGLIELTPVAAGRRLSELEFN
KPREYLDYSLVPTVYSPHPPIGVGLTEPQAIIBQYGAENKVKYSSTPAVTATVOR
OPCKMKLVGVKDEKVVGLHIGIFGVDEMIGRAVALIKMATAKADPNTVAIHTGSE
EFTVMR"
gene
complement(3218..4090)
/gene="psd"
CDS
complement(3218..4090)
/gene="psd"
/function="lipid metabolism"
/codon_start=1
/transl_table=11
/product="phosphatidylserine decarboxylase"
/protein_id="AA62138.1"
/db_xref="GI:644855"
/translation="MMTSNSYVWRLKFAQVYVMPQIYVQIAGFAKQMKITHEVYI
KAPAKKYNIMDSIAOKRQSPDYASFNEFPIRPLEKRNAPINONPTALCLPADGRISC
GHIDNLLIOAKGHFSELDLIAEDKELIVETPKNGEVTYTLSPDRHVRHMPDALT
RKMTYVPGDLFVNPFLAQHVPNLFAHNEVYICVDFEFTQVQILGATITTSIGTT
WAGYINPRHNEKVTWTYEGESAVKILKGQEMGFQIGSTVINLFOANOVRLADHLSV
NEPVRMGEIILAYKK"
misc_feature
complement(3364..3374)
/gene="psd"
/note="encodes DNA uptake recognition sequence"
complement(3780..3790)
/gene="psd"
misc_feature
/note="encodes DNA uptake recognition sequence"
3818..3825
/note="OxyR recognition sequence"
complement(4097..4100)
/note="psd"
protein_bind
complement(4126..4152)
/note="OxyR recognition sequence (Toledano, M.B. et al.,
Cell 78:897-909.)"
/bound_moiety="OxyR"
4247..4771
/gene="trpF"
4247..4771
/gene="trpF"
/codon_start=1
/transl_table=11
/product="30K protein"
/protein_id="AA62139.1"
/db_xref="GI:644856"
/translation="MOKVRLPIYVDPDKDORRLDYGVYAAANOLERLASVNVLS
AOYTLSPYVPOKILVYMKGRVQIIVDDECORCNENPQOTLECEFTYSPPVANNPOADL
PEIIEPIEFNREGESIDILIGYVEDELLIALFLVPMHSEKREVSNOBOYFGVLPBELAK
KPNFVALANKOK"
CDS
complement(4561..4926)
/note="orf121"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AA62140.1"
/db_xref="GI:644857"
/translation="LQLRGVSEFHRLLYLLIOQLLKHRYNDAHYHNEIFCEVECP
LAILNLIVFCFKLIANTANGFGFANSSGTPKTCSCADPTSQCSDECIGTSGRAKISS
SSTVPIRISPSINSIGS"
BASE COUNT 1598 a 862 c 938 g 1528 t
ORIGIN
Query Match 7.5%; Score 100.6; DB 1; Length 4926;
Best local Similarity 68.6%; Pred. No. 5.8e-14;
Matches 153; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
Oy 1 atgattaatacaaaagcgtactaacctgcacatcgagcagacgagcaagcgctt 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 224 atgattacatttaagaagagtttgatctcccaattgagggaanacgacacagtaatc 165

AUTHORS Calogero,S., Gardan,R., Glaser,P., Schweizer,J., Rapoport,G. and Debarbouille,M.
TITLE RocR, a novel regulatory protein controlling arginine utilization in *Bacillus subtilis*, belongs to the NtrC/NtrA family of transcriptional activators
JOURNAL J. Bacteriol. 176 (5), 1234-1241 (1994)
MEDLINE 94156824
REFERENCE 6 (sites)
AUTHORS Yoshida,K., Seki,S., Fujimura,M., Miwa,Y. and Fujita,Y.
TITLE Cloning and sequencing of a 36-kb region of the *Bacillus subtilis* genome between the *gnt* and *iol* operons
JOURNAL DNA Res. 2 (2), 61-69 (1995)
MEDLINE 96093926
REFERENCE 7 (sites)
AUTHORS Gardan,R., Rapoport,G. and Debarbouille,M.
TITLE Expression of the rocDEF operon involved in arginine catabolism in *Bacillus subtilis*
JOURNAL J. Mol. Biol. 249 (5), 843-856 (1995)
MEDLINE 95311309
REFERENCE 8 (bases 1 to 36200)
AUTHORS Kasahara,Y., Nakai,S., Yoshikawa,H. and Ogasawara,N.
TITLE 36kb sequence between *gntZ* and *trnY* of *B. subtilis* genome Unpublished (1997)
JOURNAL FEATURES
source location/Qualifiers
1. .36200
/organism="Bacillus subtilis"
/strain="168"
/db_xref="taxon:1423"
1. .32
/gene="gntZ"
/gene="gntZ"
/gene="gntZ"
/citation=[6]
/codon_start=3
/transl_table=11
/product="6-phosphogluconate dehydrogenase"
/protein_id="BA11267.1"
/db_xref="GI:1064781"
/translation="GVFHTNWE"
496. .1059
/gene="ahpC"
496. .1059
/gene="ahpC"
/gene="ahpC"
/note="induced by hydrogen peroxide"
/citation=[3]
/codon_start=1
/transl_table=11
/product="alkyl hydroperoxide reductase"
/protein_id="BA11268.1"
/db_xref="GI:1064782"
/translation="WSLIGEVLPPEAKAFKNGEFTDVTNEDLKQGMVFCFPADFS FVCPELEDLODYALAKELGVESVSDTHFVHKWDSSEKISKITVAMGDSQ TISRNFVDIETGLADRGFTIIDPGVIOFVEINAGGIGRDSNINVKKAQYRQ NPGVCAPKMEGGETLPFLDLVGI"
1073. .2602
/gene="ahpF"
1073. .2602
/gene="ahpF"
/note="induced by hydrogen peroxide"
/citation=[3]
/codon_start=1
/transl_table=11
/product="alkyl hydroperoxide reductase"
/protein_id="BA11269.1"
/db_xref="GI:1064783"
/translation="AVLDANIKAOINQVMOLENDIVLKVSGEDDTSKMLAVDEL ASMSKISVEKAEKELNRPSPFSYNRVGGEDVGFAGIPLGHEFTSLVALLOVSGRPK VDOKYIDOVKKISGEYHRESYISLCHNCPDVQALNNMSVLPNITTHMIDCAATKA EVESKNIMAVPTVIYNGSFGSGRMTLEITLAKMSGTDASEADPEPDVAVVGGGP AGASAAITAKRGITGVAAEREGGVDTMKSINENFISVAKTEGRLASLEBHVKEY DIDVNNIDRARLKRKDLFELENGAVLKSCTVILSTGARMRNVAVPEGEFRKNGV AYCPHCDGPLEFGKDVAVVIGGNSGIEAIIIDAGIYNHTVLEFAPELKADEVIAOKRL YSLPNVYVKAQGTKEITGDOSVNGITVYDRETGEKHELVGCVFQIOLVPTENTLE

gene complement(2712. .4151)
/gene="bg1A"
complement(2712. .4151)
/gene="bg1A"
/citation=[2]
/codon_start=1
/transl_table=11
/product="phospho-beta-glucosidase"
/protein_id="BA11270.1"
/db_xref="GI:1064784"
/translation="MGNMKDFLMGALAAHOPEGGMNOGKGPVYDVMTACAHGCP RKIYTIENEFYPNHEAIDFTHRTKEDJALRKEWGLKCLRTISGRITPKDEAP NEAGIQFVDDVDELKRGIEPIYTLSHPEMLHLAREYGGFRRNVVDFEFAEAC FTRVYDKVYVMVTFEINNMQVNNPLFLMTNSGVVYGENENAKEVYOTAHLEVAS ALAVAKKDINPEFOIGAVSHVPIPFSSNPEDVMLAEEMRORFEFPVQRYGYP SYALKEFEREGYNITEEDGDDLEIRGTVDYDNGEYSYMSYTVSKDNDNTGDVNGC LPNGVENPYITSSDWGMATDPTGIRYTLNRFYDRYQIPLPIYENFGADVTLDEDKV HDPERIOYIKSHIEALKRAVITYDGVNLTGITYPMGIIDIVSFTTGEMKKRGMITYDRO NENGSMKRYKRDSEFWKNVITNGEEL"
4733. .5449
/gene="yydK"
4733. .5449
/gene="yydK"
/codon_start=1
/transl_table=11
/protein_id="BA11271.1"
/db_xref="GI:1064785"
/translation="MRMLKYQIATEIETIYIEBHOLOGDKLPVLETLMAOFESKST ITRSLLEOKAIFOVNRSIGIFVRRKRGYISLSNGFPRKDLPEDPVATSVIELD VRKPTPEAEENINIGMEDDIYVYKRYRITNGOTLCYEEYTKSYITVYNNELYSI EHYIREGGLKIGESDLELVGQLNEEAETVIGLEGLPKLYTESIFHLNCOFPDYS KISYIEOSQFVQANSFPL"
complement(5766. .6488)
/gene="yydJ"
complement(5766. .6488)
/gene="yydJ"
/codon_start=1
/transl_table=11
/protein_id="BA11272.1"
/db_xref="GI:1064786"
/translation="MKLEFKKSISNKKVILIGAMFVFLFLGIFLLVGGIDKYSNYPE MFPSSYVATQFGLMFLPSFVIAFPIINBESNKKILFPKLIAGENITYFEPRKIAVFL ECFAPITGLIISLMYHDFSHALLLEFSAVILLOYLITGTSVLCNKLISIGYS IYVMYTVIIVAIASNKTRGFIAPFAGNTMPRIERIVLQSDMTIGSNDVLTITLV SIILINAVLRFSTRTWKGL"
complement(6509. .7138)
/gene="yydI"
complement(6509. .7138)
/gene="yydI"
/codon_start=1
/transl_table=11
/protein_id="BA11273.1"
/db_xref="GI:1064787"
/translation="MNIAVYTLKVKYKTLQDTDLHFSSGKINHYVNGKNGKYSOLAK DFLNNSKRIGRIDIRONVSLISSSSNIPNDVSKDFLHLFSKFPKAKMIDKAYLNL DNIDGKVLKINLSDGOKOKLISLLEDKNIYVDELDTNSLDKTYVIEHGFLLKXI OENPEKIIINTHRLSDLKATEGDYIIFNHOEIQYHSVDKLEIYVNE"
complement(7288. .8046)
/gene="yydH"
complement(7288. .8046)
/gene="yydH"
/codon_start=1
/transl_table=11
/protein_id="BA11274.1"
/db_xref="GI:1064788"
/translation="MKILKYEDGTYEVLYVONNVEIKDKKSGEYKNSLNSDKQLR FKMYKEVSPKRFYFLFSLPALMILNYTHILIKONGLSVYGRKMMIIEHYIYIMN IYHLEIGHISILKFGKNGDVKGLNVTYPAVYVOLNENYMLSRNKKIIVHLEGLF INVLLITLIEINQFSSSEALTMAPMFSTLMLNLTPLINSQYKILLAPLSLDEY SRKTRINMLVLTTOIIGTGAVNSVHMILYIVN"
complement(8027. .8986)

gene


```
/note="similar to hypothetical proteins from B. subtilis"
/codon_start=1
/transl_table=11
/protein_id="CAB15936.1"
/db_xref="GI:2636446"
/db_xref="SWISS-PROT:P42314"
/translation="MELIITLALGLMFAVRGFSVILFAPICALFAVLLTDPSSHL
PEFSIFEMRMAGFRLKLEPFLGALFGKVEVAGLAASIAKTIIVRLVGAARALIAI
VLMGAVLYSGVSLFVVFVAVPFAKMPQENPKRLTGTITAGLAFPTMDALPGT
POLQAVIPTSFEKTDIYAAPMGLMGAVIYLAQMILSERRKRAQASGEGFSDS
NAPAESISAAPDPKSPIRHALAFPLIIVGAVNKFTITLFWPNGEFPISGLK
EGRGRDISSAAIWSVIALVIGITITLFDMSVFAQLKEGIEGIALASMGNG
AEFGFGGIIAALFGRFKLSGISHTFTDPLVNGAVTTALAGITGSASGGGIALSM
SEQYLOAIQAYNIPPEVMHRYISMASGMDTLPHNGAVITLSCGRDPPAIIISRYFC
DHAH"
complement(3762..4645)
/gene="yxjB"
complement(3762..3788)
/gene="yxjB"
complement(3797..4645)
/gene="yxjB"
/function="unknown"
/note="similar to hypothetical proteins"
/codon_start=1
/transl_table=11
/protein_id="CAB15937.1"
/db_xref="GI:2636447"
/db_xref="SWISS-PROT:P42313"
/translation="MKRRYDFSMFRCPICDSSDAASGSKICTERHRTDLSRGV
NLTLPKRVSGCAELFEARSRLIGEGFPDPLDAIAELISHRSGHEATIIIDSGG
ECSHINALCGFYAGKAAIGTIGLSDKDIILKSKAFKIDLMVAADVARAFHDROP
VLSIFSPSNVAEFHRLKNDGMLKVRPSDYLILRLQFLYDPSPKRTSNTAAVER
FTANAHSRPRLRYVKTLDOQAIIHMLKMTPLAMSAPKDRVSLKEMKADITVDV
ILIGMK"
5261..6489
/gene="yxjA"
5261..6454
/gene="yxjA"
/function="unknown"
/note="similar to pyrimidine nucleoside transport"
/codon_start=1
/transl_table=11
/protein_id="CAB15938.1"
/db_xref="GI:2636448"
/db_xref="SWISS-PROT:P42312"
/translation="MYFLNIVGLIVNAVFLCSPQKKRIKWRPITITLVLLELTWM
FMLGTVGSMAIGKIGDEFTWLIAQASDIAIAFPMANETVGFESALPIITFIVT
FFDILTYGILPMLIDKIGWISKASRLKESFIOKMPGNTFALAVIRQOLTVL
SNRLITFGIMSSISGIIISYLMVATVETIPIINCIALITIANLNVHVE
DEDIITPPEKREKDEFSITSMYGMVYITLAMVIGYALISAVNKLGVHVG
LTNQTFALVSPFPLGLPVHDMAVYAOQDKMLATNEFVAMLDLKNKLSPHT
VAVATFLTSPNFSTGVNIGYTNISIDGKSTVIGRWMLVSGIAVSLLSAIV
GLFVW"
6462..6489
/gene="yxjA"
complement(6925..7167)
/gene="yxjA"
complement(6925..7167)
/gene="yxjA"
/function="unknown"
/codon_start=1
/transl_table=11
/protein_id="CAB15939.1"
/db_xref="GI:2636449"
/db_xref="SWISS-PROT:P42311"
/translation="MKNNMLKAAGIAVLFSVFAVAAFLKAVQAKPTVSTHTYKK
IKELTYPOVHHVGNAAFEKKTINOELKAVMDNRIRNT"
complement(7290..7306)
/gene="yxjS"
complement(7290..7602)
/gene="yxjS"
complement(7312..7602)
/gene="yxjS"
```

```
/function="unknown"
/codon_start=1
/transl_table=11
/protein_id="CAB15940.1"
/db_xref="GI:2636450"
/db_xref="SWISS-PROT:P42310"
/translation="NNRKLQTAEDENMMILVFAOKVNHDDPMELAKAYPOOK
LNESLAKRTMDDLVYRKHEEHIIPDQTVIAVLEFNTDIAQALHEAISGRK"
complement(7652..9712)
/gene="katB"
complement(7652..9712)
/gene="katB"
/EC_number="1.11.1.6"
/note="alternate gene name: katE"
/codon_start=1
```

Query Match 3.4%; Score 46; DB 1; Length 215534;
Best Local Similarity 60.3%; Pred. No. 1.7;
Matches 76; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Oy 51 gaaagccgtttacgagccgagccattaccgagtcgcgttgcgttgagcaaatatgc 110
Db 122745 GGGTCAGCGCTCGGCGTGAATTGCTTCGCGCATCGGCGCTGATACGGTGACATTTG 122686

Oy 111 cgtatgcagccctcgaatgaagcaaggaagcagtcgctgaataaagagccaaagtgc 170
Db 122685 CGGACGCGCTTTTCTTCTTGAAGTAAAGAGGCGATACCGTCAAAAAGCGAAGTGT 122626

Oy 171 gtttga 176
Db 122625 TGTGGA 122620

RESULT 23
OSJN00193/c OSJN00193 109730 bp DNA linear HTG 27-DEC-2001
LOCUS Oryza sativa chromosome 4 clone OSJNB0004G23, *** SEQUENCING IN
DEFINITION PROGRESS ***, in ordered pieces.
ACCESSION AL662992
VERSION AL662992.1 GI:17998504
KEYWORDS HTG: HTGS, PHASE2.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Euphorbiaceae; Oryzae; Oryza.
1 (sites)

REFERENCE
AUTHORS Han,B., Feng,Q., Huang,Y.C., Li,Y., Zhu,J.J., Zhao,Q., Hu,X.,
Liu,Y.L., Mu,J., Yu,Z., Chen,L., Fan,D.L., Wang,Q.J., Zhang,L.,
Lu,Y.Q., Yu,S.L., Liu,X.H., Lu,T.T., Zhang,Y.J., Lu,Y., Li,C.,
Li,T., Zhang,Y., Hu,H., Jia,P.X., Qian,Y.M., Yang,K., Zhou,B.,
Chen,Z.H., Hao,P., Zhang,L., Wu,M., Zhang,R.Q., Qian,J.P., Fu,G.,
Wang,S.Y., Ren,S.X., Lv,G., Lin,W., Gu,W.Q., Zhu,G.F., Tu,Y.F.,
Jia,J., Yin,H.F., Zhang,Y., Cai,Z., Chen,J., Kang,H., Chen,X.Y.,
Shao,C.Y., Sun,Y., Hu,Q.P., Zhang,X.L., Zhang,W., Wang,L.J.,
Ding,C.W., Sheng,H.H., Gu,J.L., Chen,S.T., Ni,L., Zhu,F.H. and
Hong,G.F.
Direct Submission
Submitted (27-DEC-2001) Han Bin, National Center for Gene Research,
Chinese Academy of Sciences, 5004 Cao Bao Road, Shanghai 200233,
CHINA. E-mail enquiries: dhan@ncgr.ac.cn. Clone requests:
dhan@ncgr.ac.cn
Oryza sativa japonica (nipponbare) genomic DNA, chromosome 4, BAC
clone: OSJNB0004G23.
Web site: http://www.ncgr.ac.cn

TITLE
JOURNAL
REMARK
COMMENT
Web site: http://www.ncgr.ac.cn
Summary Statistics
Assembly program: Phrap

FEATURES

* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers

FEATURES
source
1. location/Qualifiers
1. 24268
/organism="Streptomyces avermitilis"
/db_xref="taxon:33903"
/note="Actinobacteridae"
1. 1875
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="BAB69274.1"
/db_xref="GI:15824108"
/translation="MSAYHSPWIAVAVVIMFVVASYSRONVHAVPSGGSEYVANT
NIGPRAGITVASALVDVYLVLTAVASIASGTENIGSAIPFVEHKKSCANAAYVLLTLM
NIGVSESKLPAPITPYFVAVGTIMAMKAFRGVLDITMHPADYHITKAHOGIA
GRVLVLLRAFSGGCAALTGVEAISNGVPAFRPKSKNAASTLAAMGLAVTMCGI
IALAMATNVMAMPNDLJHNGVAVGPDYQNPVISOVAEVEFGSFLFVLAAT
ALVFLAANTVANGFPLGSLIADDRYLPRLHTRGDRLAESNGIVLLAGANTLAVI
YGADSTRLLQLYIVGVFSFTLSQSGVNRHNRHLATEKDOAKRRHVRSRALNMG
FPGVLYLVVLYMKFTHGAVALLGKCIFFYTMGAIKRHYDVAEITAAPEGSDSV
RRSRVSVLISKIHPTLALATKLMRSDTLEALSNVDPAEIKALK EEWDRKID
VPLKLVDSPEYREITRPVIEYVKGKRSKSPDAVSIIPEYVGHVY EHLHNSALRL
KGRLLTPGVMTSVPYOLESSBAKKRRARRROEMSAAGAVARGPAEERAKSSMTKG
"
1995..3296
/codon_start=1
/transl_table=11
/product="putative methyltransferase"
/protein_id="BAB69275.1"
/db_xref="GI:15824109"
/translation="MGEEVEVEIGPVAHGHGCIARTSRGOYLFEVRHALPGEVILAVT
EEBEGEYLRADAVEITLDAKSDVEIEAPCAAPRGCGGDMQHKAPGQRRLKGEVIA
BOLQRLAGLTPREAGNDGYTPABAGDKLPAGEVPAMKRVYADADGNAGIRRRSH
EPELIEHCMTAAFGVSELGIEEDWSGMSVDAIAAGSODRMVILPERPGARLVE
LDKPSVMEVEEKDGIHRVHRAVFERADGRTRYVSGSGFMVHPMAADTLVAVM
OQLPKKGMALDYCGVGLFAGALADRLDKGKAVLGIESGKRAVEDARHNLAAERV
RLEOGKEAVLPRTGITTEVDLVIDPRAGAGKKTQQLVSLGAKIAYVACDPALALA
RLGYFRDGGYKRLRAPDLFPMTHIVECVALLPEPAKGL"
complement(4334..5074)
/codon_start=1
/transl_table=11
/product="3-oxoacyl-(acyl carrier protein) reductase"
/protein_id="BAB69276.1"
/db_xref="GI:15824110"
/translation="MPTPGDGRVALVLTGASGALGAANVARYLDHRGYRLALHYNGED
QARKVAAKLSGPTLIRADVADMAATRDMDRVRAELGGISLVNTGAIIRDGLMATQ
SVDDWTRVAVANLIGTFHCRAALPDMLRGRGRIINNVSPAGLSESGRGQAYSSKA
GYLGMTRSLAECGRFNVTYNALSPGLMATALEEVPAEYATLIVARIAPFGMGTPDE
VARGTEILDADYMTGOVLSIDGMSIS"
complement(5513..6217)
/gene="cipX"
/complement(5513..6217)
/gene="cipX"
/codon_start=1
/transl_table=11
/product="CipX homolog"
/protein_id="BAB69277.1"
/db_xref="GI:15824111"
/translation="MTQITPPVYLDELITAITENHTTALDOLSGAILLADHIGEYADH
LIGHVDARHSGASWTDIGRCMGVROAKRFRVCGDAKPSNAFEKSPARSATII
AGMNAKTVGNAEIVPAHVLLGLITGEGSAVGAROGIDLDQADRTAVATLPRSDA
DPVILPYDAHAKKALEIASRVARLQODSVSTGHMLGLLEEGGAGVLAISLGLDTAA
AERFYAEADESSVAS"
complement(6353..7513)
/codon_start=1
/transl_table=11
/product="cytochrome p450"
/protein_id="BAB69278.1"
/db_xref="GI:15824112"
/translation="MLPAYETLRENGIQVELPTGKLVMMVYVPEYARIVLSDPRES
SDKTDPRPKLPNSMLKRYCAPFMINIDGPEHLKKKSINDEFSEGLARLLPRLR
VAVEKIDMDLROPTKPDVLEKALFPIAMOLEMFLGIPAAELETRDQNVKLLTGT
TTEAEREADRLNGHAEVLEKAKHIGDDMSRLIVOREKHGEVNNVELAELNLS
NAQGHNSVTMISLGVLLNLNHPDRPTLLAHPRDRTTAVDEMLRTFSVNDGTPMRL

CDS
AIEDMLIGDTLVKAGDVAVPTLPVNRDPSVCPYPHOLDIMREEPARHLAFGHGPHKC
PADRLVPSLLEIVYTTLEFERYPTLALAVEAEELTYKHSIOARGPAEMPVTW"
complement(7555..8232)
/codon_start=1
/transl_table=11
/product="cyclase"
/protein_id="BAB69279.1"
/db_xref="GI:15824113"
/translation="MPSITVGAFGSGNAVADYTKIIOEFDGEVASELPOLRRROVPSY
SDLYVMODMVDADDEAVRITSTDPGRRLMERISTGYFADNAATDMQSPVPRPE
RFYLMAPRLDDLEQTSYSAVYNTQROHRIIPRARLPASDAADITQKMTLROQIL
WRGILYIQDPAVANSMEAISBWTGDEPRLQVLDLLEIIPYPDQNSLATREYH
WAATVPA"
8414..9382
/codon_start=1
/transl_table=11
/product="aromataase"
/protein_id="BAB69280.1"
/db_xref="GI:15824114"
/translation="MTTLVRCILVSSSTGRVDHAIITIEGDVETFPDLADIARMPQFH
RSVAHETVHANENGEIVRHNAIAGDAVRAQOTRRTDPRGRRITFAHMEPPQATD
LRGMVFTELPNRTFRVELSHOVVAPTCADIGLPLDIRIAOESDCIATLADPTAERRD
LDLEVSTDPDLFTAGSYEDADYIYRADIMPERIPHYGLAMTEDVYVNVQFPDMQTK
GPDEGSTRVNSVRIQLEPNHLIVYKOISLPTMIDAHGHWRFTETPEGVIAEARHTSTI
KRSALSQGGATVPAEARRYLRLRLLSGMSMNLAKAYAEERAGP"
9449..10528
/codon_start=1
/transl_table=11
/product="o-methyltransferase"
/protein_id="BAB69281.1"
/db_xref="GI:15824115"
/translation="MSCRTGPDYTPAGSHEORTVSEGEVMAKETTPRGGVMAAADLL
TPMAVRAATRLRLDHLIAGARTTLEAAGVADDAALGRLDHLITPAGVLSGTGGA
YDITAMORLICEGPEDMRALLDIEGALGHAEISLHLITRTGGAAPPOOYGFVW
DDLSSDGRSEPTLMAKRTAFASPAVAGAIPTKTIIRVYVGGDGMILTILOSH
PDLRGTVYLDGPVPRARAKATAAGADLDRADIAAGSFPALPAGAGYLLSLTLHWD
DASARLILRCADAQTTGRVAVVDYFDQRYVOTGEDLMLGLYFGGROHTLQIAELA
GYVGLHTTSVTPACRYSVELRAVG"
complement(10495..11280)
/codon_start=1
/transl_table=11
/product="letr-family regulatory protein"
/protein_id="BAB69282.1"
/db_xref="GI:15824116"
/translation="MRPEKPARGAPSHSRROIAVAVIDIADSEGTAEASVRAIAKRL
GTAMTLYXRLTPEKEDLYTWMIDATGSEPEQDGLADTLTTLGARROGILLRHML
APLILATPIMGPNMRGNEROLAVIAGCGLGMDITIVNLTHSWSGAVQDELTERA
TANOSGLDRHAWRLMEPYLKSLETGEPHLISRVQVSEIGEADERFANGLAIIIDG
IEARFPYATGKDSRRPVGDTGGAPSAARRAKGQPTARSSTTE"
11434..12375
/codon_start=1
/transl_table=11
/product="ABC transporter I"
/protein_id="BAB69283.1"
/db_xref="GI:15824117"
/translation="MRKOYGDVVALDGLDLYVPSKTYVGLGPKGAGSTAVRIIATL
TRSSGSAVVAQFVARNPADYRKVITGLAGODASPDGDTGRENRIIGRISRG
ARRRDEILAEFGIAHAGDRLVGYVSGGMRERRIHIAISILRAPALVLEDPRTGADP
SRAEIMESVRIVADGTYVLTITPCLBEVEYELATRVGYIDKQILABGSPADLKAQIG
RHTDVYESADHVPASARLERLGRCEAGSGRIIVAAVQAPALPLIVSELDSGA
VIRNIGIRESDLDVFLMTGQTVGPPAQVETPPAVS"
12372..13157
/codon_start=1
/transl_table=11
/product="ABC transporter II"
/protein_id="BAB69284.1"
/db_xref="GI:15824118"
/translation="MNLADGWTLTORRLFRLRHPCGTLTLMMPSGFVYVFGFVFG
GAIGVGTSNVREFMFGVFMVTGTALSTAMVAVATQOAGVVERLRLSLPHTHVSVP
LGSGAGGITGAGVAVLWCGVAVGWRPHGVGRALAGTGLILLFRYTLIAWGYVIG
IYVRSARTARLARLPULSTIMVSNALVPTDMPGRLRLCENPLSSAAARARELFCN
PGAGPDAAMPULHPVTASILLMAGLLILIFVLSVRRPARAGL"
13225..14730

CDS

```
/codon_start=1
/transl_table=11
/product="putative oxidase"
/protein_id="BAB69285.1"
/db_xref="GI:15824119"
/translation="MRRREENTVLLIGTSMRVIVGGTGVGICTAFLAROCVSVH
VLEKRAOVNDHPAFGIDTRSAEILREAGIDIEDHVLVRGIMTGSTADITVPAKE
PEIGPAKIGSCGPDVDRRAALARELGQVDPDSRVAVATGGTATADGRVTADY
VTAADSVSRMRMDLIGITGTPGLSSIMHLNLFRAELTPPTSPFLNRERSGIL
LVNDETRNWFHPNGDPEPVERIQOQIDATGVSQDIEPILSLPWSSTARADR
ROGNWFLVGDAAHVPPVGGILNVGLADHNAIMAKIAAGDALDLYEERRVALE
TNRQTVLRAEHRDLNDASRAADRAVAGIMIVIVYRDSGAVVPEPISDME
NLVDGAPGSRPLRHWLPDGRSTVDLYDGLTGTADHMLDTRDGLTGTHAVEG
AAVAGIKEDGAVVVRPDGVGMRTATTRNAVSTQHLRVLDRLCRORA"
15054..16253
/codon_start=1
/transl_table=11
/product="cytochrome P450"
/protein_id="BAB69286.1"
/db_xref="GI:15824120"
/translation="MDFGADPYPEYANLRAEPEYROYLEGRIGLYGLVTRVEDYRKL
SDPRMSKDPNAPLDMQEAQKRPLEDRTGLTHLTTDAPENHRLRLVSTATARR
VEGLRAQVOHITDGLDITVPRQAEILGDFAPPLATVVICELGVKADDFRQMT
KDFRMTNTDSADRDGDAAPVGLRDLLEYLRLVDRKRODPADGLVDALIAARD
RLNEAEELLSMSLLVGGFETVNLGNGLTLRHPDOLALREPERLYDALEEML
RDGSEFETMRPRLPEIYAGTRIEKHGHLVILSLASANDGAEPAPDDPDVTRAP
AAVAREGAFHFLGAPLARLEGRITAFHGLRLRLGLALSVPEPDNRORSITVAGLEA
LPTVEDA"
```

Query Match 2.9%; Score 39.2; DB 1; Length 24268;

Best Local Similarity 51.7%; Pred. No. 72;

Matches 89; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

```
QY 326 cgtgcgaacttaagcgcggaagaagtgcgcgcgaactatccaatccggttgatga 385
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6837 CGGTGTCACGATGATCTCGCTGGCGCTCACCTGCTGACACCCGCGAGCGC 6778
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 386 ctgcgtgcgcaccgcgttcagcaaatctcgcgttcagtcgcgcgttcgcga 445
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6777 CGACGCGCTCGCCCATCCGACCGGATGACGACTGCGGTGGACAGATGCGCGCTACT 6718
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 446 tcttcgcaatgcgatgagacaacccgcgtgcgtgcgcgcacccaatgcgat 497
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6717 TCTCGGTCAACGACGACCACTGCGGCTCGCCACCGAGACATGCTCAT 6666
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 28
AP003402 187948 bp DNA linear PRI 06-NOV-2001
LOCUS Homo sapiens genomic DNA, chromosome 11q clone: RP11-661121,
DEFINITION complete sequence.

ACCESSION AP003402
VERSION AP003402.2 GI:16751506
KEYWORDS HTG.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 187948)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan
1-7-22 Suenhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URU:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT On Nov 5, 2001 this sequence version replaced gi:1336105.
FEATURES
location/Qualifiers
1..187948
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"

```
/map="11q"
/clone="RP11-661121"
BASE COUNT 55717 a 41857 c 40668 g 49706 t  
ORIGIN
```

Query Match 2.9%; Score 39.2; DB 9; Length 187948;

Best Local Similarity 51.7%; Pred. No. 79;

Matches 89; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

```
QY 414 aattcctgcctgcatgacgcgcgttcgcacatcttcgaatgcgatgacacaaatcc 473
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 29076 AATTCCTGCCAAGGATGATGCTCCTGGAGAGACCGGTGCTAGTGAACAGATCC 29017
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 474 gctgcgtgcgcaccctcaggttcattatcaagaagccgcgcgagattccaagcgagct 533
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 29016 TTTGGGACCCAGCTTTTGTGTGATTAACGAGGCTGACTTAAGTGTGAGCAAAAGCTT 28957
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 534 gtgtgattgagcgcgttcgacgaacgaacatccatgattgtaagcagct 585
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 28956 TTTGCTTTTCAAGAGGTTTTCACCCAGTAAATCTAGGGGTGGAATGAGCT 28905
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 29
SCSF2A 37245 bp DNA linear BCF 05-MAY-2000
LOCUS Streptomyces coelicolor cosmid 5F2A.
ACCESSION AL049587.2 GI:7768282
KEYWORDS ABC transporter; amino acid permease; BCF family;
carboxypeptidase; D-amino acid oxidase; ddah, dimethylarginine
dimethylammonohydrolyase; dehydrogenase; efflux protein;
endonuclease; enoyl CoA hydratase; IPP isomerase; metalloproteinase;
oxidoreductase; racemase; regulator; thiolase; transferase.

SOURCE Streptomyces coelicolor A3(2).
ORGANISM Streptomyces coelicolor A3(2).
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Streptomyces coelicolor A3(2).

REFERENCE 1 (bases 1 to 37245)
AUTHORS Redenbach,M., Kieseir,H.M., Denapate,D., Eichner,A., Cullum,J.,
Kinsahl,H. and Hopwood,D.A.
TITLE A set of ordered cosmids and a detailed genetic and physical map
for the 8 Mb Streptomyces coelicolor A3(2) chromosome
Mol. Microbiol. 21 (1), 77-96 (1996)

JOURNAL 97000351
MEDLINE
REFERENCE
AUTHORS Oliver,K. and Harris,D.
TITLE Unpublished

3 (bases 1 to 37245)
Bentley,S.D., Parkhill,J., Barrell,B.G. and Rajandream,M.A.
TITLE Direct Submission

Submitted (07-APR-1999) Streptomyces coelicolor sequencing project,
Sanger Centre, Wellcome Trust Genome Camps, Hinxton, Cambridge
CB10 1SA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.
David A. Hopwood, [3] John Innes Centre, Norwich Research Park,
Colney, Norwich, Norfolk NR4 7UH, UK
On May 10, 2000 this sequence version replaced gi:4584464.

Notes:

Streptomyces coelicolor sequencing at The Sanger Centre is funded
by the BBSRC and Beowulf Genomics
Details of S. coelicolor sequencing at the Sanger Centre are
available on the World Wide Web.
(URL: http://www.sanger.ac.uk/Projects/S.coelicolor/)

CDS are numbered using the following system eg SC7B7.0lc.SC (S.
coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary
strand).

The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
Usually the highest scoring match found by fasta -o is given for
CDS which show significant similarity to other CDS in the database.
The position of possible ribosome binding site sequences are given
where these have been used to deduce the initiation codon.

Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994) and the Frameplot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.gov.jp/>

Jun/cgi-bin/frameplot.pl. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (arg, gty, tgg or (act)) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions.

Cosmid 5F2A.

FEATURES

source

misc_feature

gene

CDS

```
1.37245
/organism="Streptomyces coelicolor A3(2)"
/db_xref="taxon:100226"
/clone="cosmid 5F2A"
1.105
/note="Nominal overlap with cosmid 4C6"
complement(83..859)
/gene="SC5F2A.01c"
/note="ddah"
complement(83..859)
/gene="SC5F2A.01c"
/note="SC5F2A.01c, ddah, dimethylarginine dimethylaminohydrolase, len: 258 aa; hydrolases asymmetrically methylated arginine with preference for (dimethylated) arginine over monomethylated arginine (experimental). Similar to TR:1038106 (EMBL:AB001915) NG,NG-dimethylarginine dimethylaminohydrolase from Homo sapiens (285 aa) fasta scores; opt: 323, z-score: 384.7, E(): 4.1e-14, (35.7% identity in 272 aa overlap)."
```

misc_feature

gene

CDS

```
misc_feature
complement(1090..1803)
/gene="SC5F2A.02c"
/note="Pfam match to entry PF00005 ABC_tran, ABC transporter, score 101.40, E-value 1.7e-26."
```

```
misc_feature
complement(1759..1782)
/gene="SC5F2A.02c"
/note="PS00017 ATP/GTP-binding site motif A (P-loop)."
complement(2002..3162)
/gene="SC5F2A.02c"
/note="Pfam match to entry PF00005 ABC_tran, ABC transporter, score 35.60, E-value 4e-08."
```

```
misc_feature
complement(3118..3141)
/gene="SC5F2A.02c"
/note="PS00017 ATP/GTP-binding site motif A (P-loop)."
complement(3245..4879)
/gene="SC5F2A.03c"
complement(3245..4879)
/gene="SC5F2A.03c"
complement(3245..4879)
/note="SC5F2A.03c, probable ABC transporter, len: 544aa; similar to many involved in antibiotic resistance eg. TR:Q54381 (EMBL:X79146) from the lincomycin-production gene cluster of Streptomyces lincolnensis 78-11 (578 aa) fasta scores; opt: 1624, z-score: 1696.3, E(): 0, (50.3% identity in 541 aa overlap). Contains two Pfam matches to entry PF00005 ABC_tran, ABC transporter, two Prosite matches to PS00017 ATP/GTP-binding site motif A (P-loop) and Prosite match to PS00211 ABC transporters family signature. Also contains a possible coiled-coil region around 250aa."
```

```
misc_feature
complement(3290..3787)
/gene="SC5F2A.03c"
/note="Pfam match to entry PF00005 ABC_tran, ABC transporter, score 122.30, E-value 9e-33"
```

```
misc_feature
complement(3743..3766)
/gene="SC5F2A.03c"
/note="PS00017 ATP/GTP-binding site motif A (P-loop)."
```

```
misc_feature
complement(4232..4783)
/gene="SC5F2A.03c"
/note="Pfam match to entry PF00005 ABC_tran, ABC transporter, score 143.80, E-value 3.1e-39"
```

```
misc_feature
complement(4403..4447)
/gene="SC5F2A.03c"
/note="PS00211 ABC transporters family signature."
```


FEATURES	SOURCE	LOCATION/Qualifiers
1. .975		/organism="Streptomyces griseus"
		/strain="45H"
		/db_xref="taxon:1911"
1. .114		
1. .975		/function="involved in the regulation of cytodifferentiation"
		/codon_start=1
		/transl_table=1
		/product="factor C protein precursor"
		/protein_id="AAC97368.1"
		/db_xref="GI:4038407"
		/translation="MPAAATSSPSRRGLRLGGGLVAASAFGLSGSGSAAAVPATK RFSLETPSHDLFRHAKLHDARVQSSFPFDYNNRLFLPAQLKSSGSPDSGKCTIOLDD SGNLTGHMYLLLGFEGHVSIGAQPVGADTYLMTEDVANSMAKGTLLARKNNKATLLSR TSSLAHQPVPAGATTEMTCALDIPVNNMRAJRLYLASGRGVIYVADIAAAYDKRPLUS DVPHPLGTGFGOYALGSGYVYQLTGNPNKIDRNPENSGSYSDVNTGALVORAFTR AGSLUTREPEPGMGIVRYTAGAEVRLFLGFASGAVAGDRRSNLFYKNVL"
115. .972		
		/product="factor C protein"
		/82 c 303 g 141 t

Query Match	2.8%	Score 38	DB 1	Length 975
Best Local Similarity	45.8%	Pred. No. 1.2e+02		
Matches 131	Conservative	0	Mismatches 155	Indels 0
			Gaps	0
QY 1026	gccgcgaacaatctccatcacgcgtacacaacctcgcattctcctgaanaaacaactctt	1085		
Db 192	GGGGTCCAGCAGTGGTTCACCTTCGACATCGTCAACGGCGGGCTTTCGTGGCCAGTT	251		
QY 1086	caagttcaacaacagccgttcaacggcgagccgcgcctgtggtccgattgtaactaga	1145		
Db 232	GAAGTCGGGCTTCGCCGACGACAGGGGGCAGCTGTGCATCACCCACGCTGACTTCGGG	311		
QY 1146	ggcgctgatgcctcttgatatactcgtcccaacctgtcttcttgcgcgattatcgttcgcga	1205		
Db 312	CAMCAAGCTGGGCGCACATGTACTCTGCTGGCTTCGGGCGACGGCGTCTCATCGGGGCCCA	371		
QY 1206	taccgacagcgcgcgcgaagcattggattgcttgtaattggaagaagaacctgcattgtg	1265		
Db 372	GCGGTCGGCGCGGACAGCTACTGTGACGGAGGTGGACGTTCACAGCAACGCGCGGTGG	431		
QY 1266	cagcttcgtctgcgcggcgaaatacgaatagcggccgcctggtgcgc	1311		
Db 432	CACCCGGCTGGCCCGTTCACTGGACACACGCGCGACGCTCAGC	477		
RESULT 32				
LOCUS	ATY14325	1694 bp	mRNA	linear
DEFINITION	Arabidopsis thaliana mRNA for mevalonate diphosphate decarboxylase			PTN 25-JUN-1999
ACCESSION	Y14325			
VERSION	Y14325.1	GI:2288886		
KEYWORDS	mevalonate diphosphate decarboxylase; MVD1 gene.			
SOURCE	thale cress.			
ORGANISM	Arabidopsis thaliana			
REFERENCE	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;			
AUTHORS	Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:			
TITLE	Rosidae: euroside II: Brassicales; Brassicaceae; Arabidopsis.			
	1 (bases 1 to 1694)			
JOURNAL	Cordier, H., Karst, F. and Berges, T.			
MEDLINE	Heterologous expression in <i>Saccharomyces cerevisiae</i> of an			
REFERENCE	Arabidopsis thaliana cDNA encoding mevalonate diphosphate			
AUTHORS	decarboxylase			
	Plant Mol. Biol. 39 (5), 953-967 (1999)			
	99273998			
	2 (bases 1 to 1694)			
	Cordier, H.			
TITLE	Direct Submission			

BASE COUNT	ORIGIN	COMMENT	FEATURES	SOURCE
480 a	352 c	375 g	487 t	
JOURNAL				
Submitted (24-JUL-1997) H. Cordier, IBMIG, 40, avenue du Recteur				
Pineau, 86022 Poitiers Cedex, FRANCE				
Related sequence: T44736.				
Location/Qualifiers				
1..1694				
/organism="Arabidopsis thaliana"				
/db_xref="taxon:3702"				
118..1356				
/gene="MVD1"				
118..1356				
/gene="MVD1"				
/EC_number="4.1.1.33"				
/codon_start=1				
/product="methylmalonate: diphosphate decarboxylase"				
/protein_id="CAA74700.1"				
/db_xref="GI:2288887"				
/db_xref="SPTREMBL:O23722"				
/translation="MAEEKVVMVTAQTPTNIAVIKYMGKDEYRILPINDISVTLDD				
PDHICLTITVAVSPSPDRDWMILWNGKEISLGSRYONCLREISRADVDVEKEGIKK				
AKKWEKHLHLHASNPPPTAAGLASSNAGACVPLAKIMNVEDPSOLATAROG				
SGSRSEVLSGGEVKMMNGKNKEDGSDVAVOLVDKDHMDLVYIIAAVSSRKETSSTS				
GMGRSVETSLDQRAKEVYVPRILQMEEALEKNDPSTFTKLCSDSNOFFAVCMDSFS				
PIPYMDTSHRIITSLVEKNRMSGTPEIATTFPAGCANVAIIANRRVAVALLGGLLIY				
CFPEKPDIMKSYVIGDTSYKKEAGLEGELQGIKDKIGSDQNGEVSFYICSPGRGKG				
PVVLQDOTQALHPTGLPK"				

Query Match	Similarity	55.8%	Pred. No. 1.4e+02	Length 1694;
Best Local	Matches	72;	Conservative	0;
Mismatches	72;	Conservative	0;	Mismatches 57;
Indels	0;	Indels	0;	Gaps 0;
OY	128	tgaagtcagaaggaagcgatgccgtccataaaagcgcaagtcgtctgttgaagcacaaga	187	
Db	71	TCGAAGCTACACAGATCTCTGTGGAATCGAAGAAAGCTTGACCTCTGTTAATGCGGAGGAGA	130	
OY	188	atccggagcgctgtgttctactgcgcgcgcttcagcgcaaaatcgcgcgcgattccgctg	247	
Db	131	AATGGGTGTGTATGTGTGACGCGCCAGACTCCACAGCAATATGCCCGTGAATAGTATTTGGG	190	
OY	248	aaaagcgcg 256		
Db	191	GAACAGAGAG 199		
RESULT	33			
LOCUS	AE004619			
DEFINITION	Pseudomonas aeruginosa PA01, section 180 of 529 of the complete genome.	11351 bp	DNA	linear
ACCESSION	AE004619	AE004091		
VERSION	AE004619.1	GI:9947920		
KEYWORDS				
SOURCE				
ORGANISM	Pseudomonas aeruginosa. Pseudomonas aeruginosa Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; Pseudomonas.			
REFERENCE	1 (bases 1 to 11351) Stover,C.K., Pham,X.O., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S., Huftagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Goltzy,L., Tolentino,E., Westbrock-Adman,S., Yuan,Y., Brody,L.L., Coulter,S.N., Folger,K.R., Kas,A., Larbig,K., Lim,R., Smith,K., Spencer,D., Wong,G.K., Wu,Z. and Paulsen,I.T., Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen			
TITLE	Nature 406 (6799), 959-964 (2000)			
JOURNAL	2 (bases 1 to 11351)			
MEDLINE	Stover,C.K., Pham,X.O., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S.L., Huftagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Goltzy,L., Tolentino,E., Westbrock-Adman,S., Yuan,Y., Brody,L.L., Coulter,S.N., Folger,K., Kas,A., Larbig,K., Lim,R., Smith,K., Spencer,D., Wong,G.K., Wu,Z. and Paulsen,I.T., Westbrock-Adman,S., Yuan,Y., Brody,L.L., Coulter,S.N.,			
REFERENCE				
AUTHORS				

VERSION AL646061.1 GI:17427781
KEYWORDS
SOURCE Ralstonia solanacearum.
ORGANISM Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
Ralstonia.
REFERENCE 1 (bases 1 to 198050)
AUTHORS Salanoubat,M., Genin,S., Artiguenave,F., Gouzy,J., Mangenot,S.,
Arlat,M., Billault,A., Brottier,P., Camus,J.C., Cattolico,L.,
Chandler,M., Choise,N., Claudel-Rehard,C., Cunne,S., Demange,N.,
Gaspin,C., Lavie,M., Moisan,A., Robert,C., Saurin,W., Schlex,T.,
Siguer,P., Thebaud,P., Whalen,M., Wincker,P., Levy,M.,
Weissenbach,J. and Boucher,C.A.
TITLE Genome sequence of the plant pathogen Ralstonia solanacearum
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 198050)
AUTHORS Boucher,C.A.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2001) Genoscope and CNRS UMR-8030, 2 rue Gaston
Cremieux, CP5706, 91057 Evry Cedex, France, Laboratoire de Biologie
Moléculaire des Interactions Plantes-Microorganismes INRA-CNRS,
BP27, 31326 Castanet-Tolosan Cedex, France, Fondation Jean
Dausset-CEPH, 27 rue Juliette Dodu, 75010 Paris, France, LMG CNRS
URGV, 2 rue Gaston Cremieux, F 31062 Toulouse Cedex 4, Genoscope and INRA
Laboratoire de Biometrie et Intelligence Artificielle INRA, BP27,
F31326 Castanet-Tolosan Cedex, Laboratoire de Genetique Cellulaire
INRA, BP27, F31326 Castanet-Tolosan Cedex
Christian Boucher@toulouse.inra.fr
http://sequence.toulouse.inra.fr/R.solanacearum.html
FEATURES
source Location|Qualifiers
1..198050
/organism="Ralstonia solanacearum"
/strain="GMI1000"
/db_xref="taxon:305"
151..642
/gene="RSC0770"
/note="RS05084"
151..642
/gene="RSC0770"
/EC_number="1.-.-.-"
/function="small molecule metabolism"
/note="Product confidence : putative
Gene name confidence : hypothetical
Predicted by Codon usage
Predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="PUTATIVE TRANSMEMBRANE DEHYDROGENASE (SMALL
SUBUNIT) OXIDOREDUCTASE PROTEIN"
/protein_id="CAD14472.1"
/db_xref="GI:17427782"
/translation="MHTRPPSSGDGDPAGLTRROWLSALAAFAASLVGSVYLLXAVAO
APGPIIDAFMGISQALTRPSALDRAVAGRIALLAIOASGPAQLRPLAOSIAAGTIS
GIQGETAIRLIQAWTGYVDGIVYIEQALMGVSDIIVITRYICENRGEVMAEPEVE
KOS"
642..2261
/gene="RSC0771"
/note="RS05083"
642..2261
/gene="RSC0771"
/EC_number="1.-.-.-"
/function="small molecule metabolism"
/note="Product confidence : putative
Gene name confidence : hypothetical
Predicted by Codon usage
Predicted by Homology
Predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="PUTATIVE TRANSMEMBRANE DEHYDROGENASE (LARGE

SUBUNIT) OXIDOREDUCTASE PROTEIN"
/protein_id="CAD14473.1"
/db_xref="GI:17427783"
/translation="MAETROAQBOADIIVVSSGVAALVAYELARAGKSVLMRAGPRL
PMEIYERPRRNOADKDDPMAPYSTWMAHPEPGPNNTIILGKHFDSOYIRAVGG
ITWHAASWTRRLPNDFKLRSYGIRADMPLOYDDLERRYGAEALGVGPNDEIG
5FRSQDYPAPLPLESTNKERTVYDALNAHDPALVAVPEPARNSRPYDGRPTCCGNNC
MPCIFGAMNGIVHVEKEAOAGARLIENAVVYKLEVGAKRIVAAHYKDPKGVDRHY
EKGWVLANGIETPKMLMSTGADPKKGVNSDMWGRMLMHPGVGSFVADKILM
PGRQDEMVLIGFRDGPRAAPQAGKHLISNTRLEOETTRIFKAGKILKPAELAR
IRDOARVQFQSPFHEILPPEMRIVPSATGTDALGPETRIYRDDYKRSAYTHR
EYVATPAVIGTSDVQFHDFPARNNHITATIGSPKDSVVDKCDRTDHPNLFIS
SATMTVGVVNTLTIATLALRIADQLKEA"
2263..3552
/gene="RSC0772"
/note="RS05082"
2263..3552
/gene="RSC0772"
/EC_number="1.-.-.-"
/function="small molecule metabolism; energy transfer;
electron transport"
/note="Product confidence : putative
Gene name confidence : hypothetical
Predicted by Codon usage
Predicted by Homology
Predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="PUTATIVE OXIDOREDUCTASE DEHYDROGENASE
(CYTOCHROME C SUBUNIT) SIGMAL PEPTIDE PROTEIN"
/protein_id="CAD14474.1"
/db_xref="GI:17427784"
/translation="WRRKAASFILLSVLGAARADADAPADAQIKRGEYLIATAGD
CVACHTAGGKPFAGGELPLOTIMLETISSNTTDPKTCIGKRYEDFERAARQGVSD
DGENLYPAMPYPSYKAIINDADMDRLDYAFERGVAAVOVAPPGTIRMPIMPMKILM
NLAFIKRAPYLPKGDMPPTMNRGAYLAQGLAGVACGLKPRGAMOEKAMDEKCYLA
GSTLAGMAYNTSDPVSGIGGKREQIYOYLRISVIGTCGAAPKMGAVOHSYSRM
TERDVRAIAEYLRTPYPAVSTGAQRARHRMGTPATDYVILRGKFTITTDIAALYIGNC
ATCQADGKGTTPDCTYPPPLNSAVAGARHGRDRLVOMLHGERAGDRHIGMPARGQ
LSDAQLALANVYTRQEDPATPALTDIELAKRMPQ"
complement(3559..5781)
/gene="RSC0773"
/note="RS05081"
complement(3559..5781)
/gene="RSC0773"
/function="miscellaneous; hypothetical/global homology"
/note="Product confidence : probable
Gene name confidence : hypothetical
Predicted by Codon usage
Predicted by Homology
Predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="PROBABLE TRANSMEMBRANE PROTEIN"
/protein_id="CAD14475.1"
/db_xref="GI:17427785"
/translation="MRTSARVILGIGTPIYVAAIVFVLTPTMNRKPYLNDHSQ
AVGPRFINDGLATVWKPPGESGWRVYVWPRILASDIAIGNDMARQPHVAVYORL
TFVLEALPLLAHRTIVPSVTLDTPMIALERDSQNNMTFHRGADSDSPSMKLELR
IARSKGVQDDEVKRIHLATLDTVNOALYTAASGAIIRGSPAPASAGAASATD
PAAPDQOYDGIAMTVKTYNNNAIRSGRAGGVRLQDARRPYPLADVYVGRKTRID
LACTGNPSLTFLDLRLHLGASMAHLYPTGVVLRDPPFRGRILGRLRGSS
WREOFTGVRGSGDIDGTLTPARPEGQRLQGLGVSHOLEADLAPLITGAOSNAS
KQKRDAPVRPAKVLVEPFRTRDMRNAIDAVKETERIVRADLPIDHLYTHIKQ
DAVTLIDPLNEVAGGTLTSLNRKDGSAWQASAAALAAHRFKTROLFPNTESMHASI
GEINGARLSATGNSVALILGASNGELKILVEQGVASKFILLEAGLWVGNTLTKRG
DKQVSTNCAAGDAVSDGLAQARTFVVDQDAVDVYGATSEKHSKESITTHPSKGL
RVFSLRTPLYKGVKHPDVSNPVAVALRAGGVVLAFAAPVAAYVPIYELQAPDS
PCSKLADYKQRTAPRPPCKVYVRSQGTGRADSGPKDPAAPARQPALPLRDPITTAG
5883..6785

[illegible]

RBS	FLRVYAGAGRTGDAAVAVHSHON" complement(3028. .3031) /gene="gra-orf9" 3152. .3226 /gene="tRNA-ala" /product="tRNA-ala" 3152. .3226 /gene="tRNA-ala" 3232. .3267 /gene="tRNA-ala" 3232. .3267 /gene="tRNA ala" 3232. .3267
gene	
terminator	
protein_bind	/gene="tRNA ala" /note="putative terminator (stem 16 bp, loop 4 bp)" complement(3290. .3346) /gene="gra-orf8" /note="5 CGA repeats" /bound_moiety="SARP" complement(3290. .3346) /gene="gra-orf8" 3454. .3457 /gene="gra-orf9" 3454. .4313 /gene="gra-orf9" 3462. .4313 /gene="gra-orf9" /codon_start=1 /transl_table=11 /product="pathway-specific transcriptional activator" /protein_id="CA09630.1" /db_xref="GI:4218541"
gene	
CDS	/transl_table=11 /translation="MDVGIIIGPLAVLRNGSIVPSACKPRLALLAIRERVTVPT LMEIETMGAIPIKSAATYLOTYILHLRRKVTALNAGOPGANADRGAGAGDAKSVIV STCEGFGYIALIRPEDESDIRFQESLAARGSVALLEGADARASAMLERALGIRGTALV VPGVGLAEVEYIGMEQRTYRLEQRIEADLLIGRASLLATRLVLRHPNENLAAQO LMTALYRSGVWRALLEVQELRRLTLEELGVEPSRQLLHRAVLSEGLDPRHAGL ELMRFGG"
terminator	4381. .4413 /gene="gra-orf9/gra-orf10" /note="putative bidirectional terminator (stem 12 bp, loop 9 bp)" 4381. .4413 /gene="gra-orf9/gra-orf10" complement(join(4418. .5038,5051. .5140)) /gene="gra-orf10" complement(4418. .5026) /gene="gra-orf10" /codon_start=1 /transl_table=11 /product="putative two-component response regulator" /protein_id="CA09631.1" /db_xref="GI:4218542" /translation="MIRILIAEDMEMVARGALVALLGLEDDLEVAOYERGDKITPEAL KVRPVAVIIDIGLEGIDIGIKAAALLAEELPSCRTILITLSRPGVLRALDAKAGTTE PKDSAPOLAAVRPIAAGHRADIPOLALAMDNSESPITREMEVLRITRDEGDAEP IAARHETGTVRYRTIVYTKLNARNVDVARTAHESGWL"
RBS	complement(5038. .5038) /gene="gra-orf10" complement(5051. .5140) /gene="gra-orf10" /note="4 CC/AA repeats" /bound_moiety="SARP" complement(join(5125. .6345,6567. .6602)) /gene="gra-orf11" complement(5125. .6345) /gene="gra-orf11" /codon_start=1 /transl_table=11 /product="putative two-component sensory kinase" /protein_id="CA09632.1" /db_xref="GI:4218543"
gene	
CDS	/translation="MNSREQQFDELDSSATLSHEDSLPAPRLARTVYLLAANAGTISI TVVAVIAGAGSPACTGVGTALLAVPELQVAHSPQAROSPVRWRVTTTLQLVNCFE VPLAFENHAKMGAGFLAGSVLLLPGRPAAPLYGATGASFPAYSAAGHGTLLDSVYVT

protein_bind
gene
terminator
CDS

LEATLVGLVGEVSGSLTGLVTRLIHARADRLAIVAGGERFPERSDLIHDLGLVLSLA
TIKTELVAVRIPDMDEPARDDTSTLIARSLDIDKRVKAGVAGSDMSLRREAAATBSM
LEAGCTIEPAALIDIGLPESQIDITYLATYLRGGYNALRHSHVNHCTIRCKEKDSYVTL
SLRANGVPEEGEGVGDARDARRADDPGCTGLNLTJRLTRVGGSLTLESDNGMFRRA
TARAVHVVHVTTR
complement(6567) .6602)
/gene="gra-orf11"
/note="2 CC/GA repeats"
/bound_moiety="SARP"
complement(6688) .8981)
/gene="gra-orf12"
complement(6688) .6711)
/gene="gra-orf12"
/note="putative terminator (stem 8 bp, loop 8 bp) "
complement(6716) .8971)
/gene="gra-orf12"
/function="unknown"
/codon_start=1
/transl_table=11
/protein_id="CA09633.1"
/db_xref="GI:4218544"
/translation="MTGEAVYLVSGRGVAVRGAPAEALPILDTGRTPEVLEAREAR
AADARPAPRYRRLALIELMSAGLIRAPDDAAPGTCGAPPAHDSAAEKFMLOLACT
TDPATYVRLVLDGPGVEEEDAPACASGLDVRPGQASDAEVTLYLTGIDYDPLT
ADVDLQAAAGRPWTIARPGGPFAAMTGVLARDDGPGMCHAEPLRLQSGEAPLHRA
LIGCTSPVARPPASLASRAASLIHTAVLETKAVLWAGVCPPTDNYHTVDTALDSARHP
VARRKQSCSGAGLVRAVRAPFPYRPPAPPAADGCGNHRALSREVLERTGHLGVY
TGVAVERDSDPSPAFTHTSLGSHNLARAPTLAGHSAGLKSGLSGSGKGLTSEERKVAS
LCEAEFEDSSGSHMDGDEPVRSFGFAEAAVHPVTLYRDPRLRSDHANNASGSPQY
YVGRFEDRSDAPDWTPTVMSLTGVSQRLIPTSLTYFDAPAPLRDSSGNAAGSSPDD
AFVQGFLELVEDRDAALMWYNRTAPACDLDAPFASGAWMEPLRAGYRSIGREITMAL
DUTSGFGLIYVAALSRLTGDGPEVDYLFEGFANLPRILAVRRLIEEMOLLPAVVADPAR
DGSQGTGVLVLAISMMWRATYVAGOPHLLIPDQRPASRPGSGTAYRARDLSIDVAAALAA
LADRGLEYVVLVDQTRPDIGLPVYKVVYPGKRWPPRABGRFLFDVYVALGRLAEPTR
YERLPLELVY

RBS
gene
CDS

complement(8978) .8981)
/gene="gra-orf12"
complement(9059) .9245)
/gene="gra-orf13"
complement(9059) .9232)
/gene="gra-orf13"
/function="unknown"
/codon_start=1
/transl_table=11

Query Match	2.8%:	Score 37.4:	DB 1:	Length 39250:
Best Local Similarity	55.9%:	Pred. No. 2e+02:		
Matches 71:	Conservative	0:	Mismatches 56:	Indels 0:
				Gaps 0:
QY 1090	ttcaacacagcgcgtcaaacgagcgagccgagccatggtgccgagtgtactacgagcgc	1149		
Db 13431	TTTACCGCGCGCGGTCACAGAGCGCCGCTCCGGCGATGACGAGCGCGCGCGGACACTG	13490		
QY 1150	gtgtagcccttgatatactctgcccacccctgtccttgtagcgagcttaactgtgcgagatacc	1209		
Db 13491	GAGATCACCGACGCGCATCTCACTCCTGTGGAGCGCGCGACCGGGTGTCTCCGACGAG	13550		
QY 1210	gacagcg 1216			
Db 13551	TACACCG 13557			
RESULT 43				
AF429315/c				
LOCUS	AF429315	125020 bp	DNA	linear PRI 18-JAN-2002
DEFINITION	Homo sapiens junctophilin 3 (JPH3) gene, partial cds.			
ACCESSION	AF429315			
VERSION	AF429315.1	GI:17646244		
KEYWORDS				
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2002, 06:19:59 ; Search time 27321.5 Seconds
(without alignments)
1029.420 Million cell updates/sec

Title: US-09-303-518D-127
1344
1 atgattataaatcaataaaagg.....cctltgagaagaagactga 1344

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :
1: GenEmbl:*
2: gb_ba:*
3: gb_htg:*
4: gb_in:*
5: gb_om:*
6: gb_ov:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_mu:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_inv:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No. Query Match Length DB ID Description

SUMMARIES

1	1327	98.7	1344	6	A96094	A96094 Sequence 12
2	1327	98.7	311321	1	NMA332491	AL162754 Neisseria
3	1229.4	91.5	1344	6	A96092	AL162754 Neisseria
4	1229.4	91.5	12273	1	AE002412	AE002412 Neisseria
5	1229.4	91.5	349980	6	AX044029	AX044029 Sequence
6	1165.4	86.7	1344	6	A96098	A96098 Sequence 13
7	562	41.8	1558	1	APU24492	U24492 Actinobacil
8	554.2	41.2	10563	1	AE006171	AE006171 Pasteurel
9	530	39.4	13299	1	U32702	U32702 Haemophilus
10	506.2	37.7	6048	1	AF1165980	AF1165980 Vibrio ha
11	503	37.4	5916	1	AF117331	AF117331 Vibrio ch
12	503	37.4	13382	1	AE004300	AE004300 Vibrio ch
13	477.4	35.5	5078	1	VANORHOT	Z37111 V. alginolyt
14	475.8	35.4	6908	1	AB008030	AB008030 Vibrio al
15	473	35.2	16903	1	AE004724	AE004724 Pseudomon
16	431.6	32.1	474	6	A96090	A96090 Sequence 12
17	333.6	24.8	96109	6	AX067460	AX067460 Sequence
18	122.2	9.1	1356	1	AF144076	AF144076 Porphyrom
19	101.4	7.5	4926	1	AF144076	U20229 Haemophilus
20	45.6	3.4	125020	9	AF429315	AF429315 Homo sapi
21	44.6	3.3	36200	1	BACGNRZA	D78193 Bacillus su
22	44.6	3.3	215534	1	BSUB0021	Z99124 Bacillus su
23	41.2	3.1	1641	8	AB015974	AB015974 Pseudomon
24	41.2	3.1	628	1	AF316898	AF316898 Haemolea
25	40.6	3.0	11375	1	AE002001	AE002001 Deinococc
26	39.2	2.9	340806	3	NMA122491	AL162752 Neisseria
27	38.2	2.8	2557	3	AY071018	AY071018 Drosophil
28	38.2	2.8	65140	6	AX211705	AX211705 Sequence
29	38.2	2.8	89791	3	AC005268	AC005268 Drosophil
30	38.2	2.8	101085	2	AC018341	AC018341 Drosophil
31	38.2	2.8	123580	1	AF263932	AF263932 Streptomy
32	38.2	2.8	125401	6	AX211739	AX211739 Sequence
33	38.2	2.8	164035	3	AC007888	AC007888 Drosophil
34	38.2	2.8	305020	3	AE003452	AE003452 Mus muscu
35	38	2.8	69283	2	AC110159	AC110159 Homo sapi
36	38	2.8	125020	9	AF429315	AF429315 Homo sapi
37	37.2	2.8	9743	14	AF435865	AF435865 Rubella v
38	37.2	2.8	9743	14	AF435866	AF435866 Rubella v
39	37.2	2.8	9762	6	AX009468	AX009468 Sequence
40	37.2	2.8	9762	14	AF188704	AF188704 Rubella v
41	37.2	2.8	9762	14	RUBNOPR	L78917 Rubella v
42	37.2	2.8	112721	8	AC016780	AC016780 Genomic S
43	37.2	2.8	142737	8	AC027658	AC027658 Oryza sat
44	37.2	2.8	329861	1	NMA522491	AL162756 Neisseria
45	37	2.8	763	6	AX019436	AX019436 Sequence
46	37	2.8	1598	6	AX019435	AX019435 Sequence
47	37	2.8	2042	1	AF031148	AF031148 Methyloco
48	37	2.8	9807	1	AE009707	AE009707 Brucella
49	36.8	2.7	5931	1	BCU97042	U97042 Burchholderi
50	36.8	2.7	10695	1	AE007265	AE007265 Sinorhizo
51	36.6	2.7	9762	14	AB047329	AB047329 Rubella v
52	36.6	2.7	9762	14	AB047330	AB047330 Rubella v
53	36.6	2.7	10029	1	AE008364	AE008364 Agrobacte
54	36.6	2.7	10051	1	AE009249	AE009249 Agrobacte
55	36.6	2.7	10486	1	AE001656	AE001656 Chlamydia
56	36.6	2.7	15680	1	AE002164	AE002164 Chlamydia
57	36.6	2.7	42348	1	SC27G11	ALJ311763 Streptomy
58	36.6	2.7	65979	8	AF254799	AF254799 Streptomy
59	36.6	2.7	189050	1	AL646066	AL646066 Ralstonia
60	36.6	2.7	300550	1	AP002547	AP002547 Chlamydia
61	36.4	2.7	1325	6	AX168124	AX168124 Sequence
62	36.4	2.7	38634	1	SC28G5	AL359152 Streptomy
63	36.4	2.7	188050	1	AL646072	AL646072 Ralstonia
64	36.2	2.7	4081	14	AF274291	AF274291 Culex nig
65	36.2	2.7	4623	1	AF419157	AF419157 Aeromonas
66	36.2	2.7	11525	1	AE005110	AE005110 Halobacte
67	36.2	2.7	108252	14	AF403738	AF403738 Culex nig
68	36	2.7	762	6	AX019439	AX019439 Sequence
69	36	2.7	3230	9	AK023114	AK023114 Homo sapi
70	36	2.7	10944	1	AE002563	AE002563 Neisseria
71	36	2.7	11162	1	AE005655	AE005655 Escherich
72	36	2.7	32039	1	SC7A1	AL034447 Streptomy
73	36	2.7	172325	6	AX044035	AX044035 Sequence


```

Db      1261 TTGTGACACTTCGTCTGCCCGGCAATACGATANGCCCGCTGTTCGTAAGTGCTG 1320
QY      1321 gaacacnttgaagaaggaagctga 1344
        |||||||
Db      1321 GAACACNTTGAAGAGAGGCTGCA 1344

RESULT  2
NMA32491/c 311321 bp DNA linear BCT 04-DEC-2000
LOCUS      NMA32491
DEFINITION Neisseria meningitidis serogroup A strain Z491 complete genome;
VERSION    A1162754 AL157959
KEYWORDS   A1162754.2 GI:7379424
SOURCE     Neisseria meningitidis Z2491.
            Bacteria: Proteobacteria; beta subdivision; Neisseriaceae;
            Neisseria
REFERENCE  1 (bases 1 to 311321)
            Parkhill,J., Achman,M., James,K.D., Bentley,S.D., Churcher,C.,
            Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
            Davies,R.M., Davis,P., Devlin,K., Feltham,T., Hamlin,N.,
            Holtroyd,S., Jagers,K., Leather,S., Moule,S., Mungall,K.,
            Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
            Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
            Complete DNA sequence of a serogroup A strain of Neisseria
            meningitidis Z2491
            Nature 404 (6777), 502-506 (2000)
            20222556
TITLE      2 (bases 1 to 311321)
            Parkhill,J
JOURNAL    Direct Submission
MEDLINE    Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
AUTHORS    Sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT    Notes:
            Details of N. meningitidis sequencing at the Sanger Centre are
            available on the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES   source
            1. 311321
               /organism="Neisseria meningitidis Z2491"
               /strain="Z2491"
               /db_xref="taxon:122587"
               /note="serogroup: A"
               115..120
               /gene="NMA0698"
               129..587
               /gene="NMA0698"
               129..587
               /note="NMA0698", possible periplasmic protein, len: 152 aa;
               unknown, contains a probable N-terminal signal sequence
               and lies within a region of unusually low GC content"
               /codon_start=1
               /transl_table=11
               /product="putative periplasmic protein".
               /protein_id="CA883985.1"
               /db_xref="GI:7379425"
               /db_xref="SPTREMBL:O9J0Q2"
               /translation="MTKRLILFLFEFVLAIALNLEFFSSDIESFGNYOFERYVD
               KGPANVILVKGDNCGNPKIISGLVLEYKEDNITYFSYIDGGFASDSCYKPEI
               LVGKITNKHHIINISMENKNFLSEDKIMKGTNWLADPKNKCNIQTLD"
               742..1077
               /gene="NMA0699"
               742..1077
               /gene="NMA0699"
               /note="NMA0699", pseudogene, probable ABC transporter
               protein, len: 336 bp; similar to C-termini of many ABC
               transporters e.g. TR:046973 (EMBL:U47048), mtfb,
               Escherichia coli microcin transport protein (707 aa),
               fasta scores: E(): 2.6e-19, 53.2% identity in 111 aa
               overlap. Contains PS00211 ABC transporters family

```

```

signature. NMA0686 may be the remainder of this
pseudogene"
/codon_start=1
/pseudo
/transl_table=11
/product="putative ABC transporter protein (pseudogene)"
868..912
/gene="NMA0699"
/note="PS00211 ABC transporters family signature"
complement(1101..2327)
/gene="NMA0700"
complement(1101..2327)
/gene="NMA0700"
complement(1101..2327)
/note="NMA0700", possible ribonuclease BN, len: 408 aa;
similar to SW:RBN_ECOLI (EMBL:L19201), rbn, Escherichia
coli ribonuclease BN (EC 3.1.-.-), fasta scores: E():
2.3e-28, 34.3% identity in 268 aa overlap, longer than rbn
at the C-terminus. Also similar to TR:085449
(EMBL:AF067083) Vitreoscilla sp. hypothetical protein (376
aa), fasta scores: E(): 0, 45.0% identity in 333 aa
overlap. Contains a region similar to NMA0677, fasta
scores: E(): 4.3e-08, 64.4% identity in 45 aa overlap.
Contains hydrophobic, possible membrane-spanning regions,
as does rbn
/codon_start=1
/transl_table=11
/product="putative ribonuclease BN"
/protein_id="CA83987.1"
/db_xref="GI:7379426"
/db_xref="SPTREMBL:O9JVT9"
/translation="MTFLQRLGLADNKCICAFMVFYRDEFERYPOAASMTFTLL
ALVPILTVAVASIFPVDFRMSDSFVSFVNDITVPOGADWFDYINARREQNRRLTA
IGSVMLVVTSLMIRITDNTFNRIWRVNSQRPMMQFLVYVALLTFGPISLGVGISEM
VGSVODALASGAPQMSGALRTAATLTFPMTLLMGLYRVPNRFVPAQAFVQALNTA
RCLTPARSRFTVMVGMNPDGYSRISYFAFVAPFLMLNLMTLVLGAVTSSLSYQO
GEAFRRGSDSRGRFDVIAKLILLIDPAACKEGRALPQFERRRHINMGYDELLEKLA
RHGYIYSRGQVNLKTAGDSIELNELFKLFVIRPLVERDHVQADAVAMPCLOQTIN
MTLAEFDAQAKROOQS"
1376..1385
/note="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
complement(2417..2426)
/note="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
2444..2453
/note="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
2486..2743
/gene="NMA0701"
2486..2743
/gene="NMA0701"
/note="NMA0701", possible pseudogene, len: 258 bp; shows
weak similarity to part of SW:WRHA_ECOLI (EMBL:M99166),
wrba, Escherichia coli T1P repressor binding protein (197
aa), fasta scores: E(): 0.064, 29.3% identity in 82 aa
overlap. Also similar to part of TR:085450 (EMBL:AF067083)
Vitreoscilla sp. T1P repressor binding protein (fragment)
(124 aa), fasta scores: E(): 6.6e-14, 56.6% identity in 83
aa overlap"
/codon_start=1
/pseudo
/transl_table=11
/product="hypothetical protein NMA0701 (pseudogene)"
complement(2755..2764)
/note="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
2778..2787
/note="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
2861..2864
RBS      2861..2864
gene      2871..3530
/gene="NMA0702"
2871..3530
CDS

```

/gene="NMA0702", len: 219 aa; similar to TR:052925
(EMBL:Z50189), exsB, Rhizobium meliloti putative regulator
of succinoglycan biosynthesis (not a transcriptional
regulator) (234 aa), fasta scores: E(): 3.1e-15, 32.78
identity in 205 aa overlap. Also similar to many
bacterial hypothetical proteins e.g. SW:YBAX_HAEIN
(EMBL:U32798), H1191, Haemophilus influenzae hypothetical
protein (196 aa), fasta scores: E(): 0, 79.28 identity in
173 aa overlap
/codon_start=1
/transl_table=11
/product="hypothetical protein NMA0702"
/protein_id="CA863989.1"
/db_xref="GI:7379427"
/db_xref="SPTREMBL:Q9JVT8"
/translation="MSNKKALVIFSGGSDSTCLIQAIOTYGRNVOALTFQYGRRA
VELERARWIAADLVKQVQVLDLSIMROITHMLMDPTALETRENGVPMFVGRRA
FLIYAIVAKGGIRHIIAGVETDPSGPDRCRDVFKSMNYTLNAMPYDRIHPL
MYLKAQTALADEKALDYIREQHTCTNGIVGGRECPSCILRERGLAEHLESKKA
V"
misc_feature
3522..3531
/note="Core DNA uptake sequence: gccgtctgaa"
RBS
3547..3550
/label=DUS
3558..4076
/gene="NMA0703"
3558..4076
/gene="NMA0703"
/note="NMA0703", len: 172 aa; unknown"
/codon_start=1
/transl_table=11
/product="hypothetical protein NMA0703"
/protein_id="CA863990.1"
/db_xref="GI:7379428"
/db_xref="SPTREMBL:Q9JVT7"
/translation="MPKLMFTYLGNGAGSNIEVDIQPACDDYEAVALPKAMFG
DADKIHIDMOVEMADVDYDIASEPTKPMSENPRLYFANVGSRAGQLEAFAF
GLFAATPAEAKOKALOTLLTDSYVOHDKNDLVNLLADRIGNFHLRPNPBGK
PAEIGFOGYLPI"
3782..3791
/gene="NMA0703"
/note="Core DNA uptake sequence: gccgtctgaa"
/label=DUS
4083..4505
/gene="NMA0704"
4083..4502
/gene="NMA0704"
/note="pfam match to entry PF01242 PRPS, 6-pyruvoyl
tetrahydropterin synthase, score 10.80, E-value 1.7e-06"
4083..4505
/gene="NMA0704"
/note="NMA0704", len: 140 aa; similar to many hypothetical
proteins e.g. SW:YB90_HAEIN (EMBL:U32798), H1190,
Haemophilus influenzae hypothetical protein (141 aa),
fasta scores: E(): 5.1e-32, 58.08 identity in 138 aa
overlap. Shows very weak similarity to eukaryotic
6-pyruvoyl-tetrahydropterin synthases e.g. SW:PRPS_RAT
(EMBL:M77850), Pts, Rattus norvegicus 6-pyruvoyl
tetrahydropterin synthase precursor (EC 4.6.1.10) (144
aa), fasta scores: E(): 0.0062, 27.38 identity in 154 aa
overlap. Contains pfam match to entry PF01242 PRPS,
6-pyruvoyl tetrahydropterin synthase"
Query Match 98.7%; Score 1327; DB 1; Length 311321;
Best Local Similarity 98.7%; Pred. No. 1e-312;
Matches 1327; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 1 atgattaaatacaaaaggcttaaacctcgccatcgccgagcagacgagagcaatatt 60
Db 55476 ATGATTAAATCAAAAAGGCTTAACCTCGCCATCGCGGCGAGACCGGAGCAAGTATT 55417
QY 61 tatgacgagccgcctacatcagaagtcgcttgctggcgaagaatatcgcgatgagc 120

|||||
Db 55416 TATGACGCGCCCGCATTTACGAAGTCGCTGCTGGCAAGAAATATCGCGGATTCGC 55357
QY 121 ccttngatgaatcaagaagcgcgatgcgttcaaaaaagccaaagcgctgtttgaagac 180
Db 55356 CCCCTGATGAAGCAAGCAAGGAGCGATGCGTCAAAAAGCCCAAGTGTGTTGAAGAC 55297
QY 181 aaaaagatccggcgctgtgttttaacccgcccngtttcaagcaaaatcgcgcctacat 240
Db 55296 AAAAAGAAATCCGGCGGTGTTTACCGCGCCGGTTTCAGGCAAAATATCGCGCATCAT 55237
QY 241 cggcgagcaaaagcgcgacttcaagtcgctgtgtattccgtttgaagcagcagaaatc 300
Db 55236 CCGGCGCAAAAAGCGCGTATTCATGTCGTGTGATTCCTGTGAAGCAACAGAGAAATC 55177
QY 301 gacttcaacgctacgcgcccgaagcgttgcgaacttaagcgcgcgaagaantnmgc 360
Db 55176 GAGTTCGACACGTATACGCGCCGAAGCGTTGGCAAACTTAAGCGCGGAAGATGCGCCGC 55117
QY 361 aatctgattcaatccggtttgtgactgcgtgcgttccgttccagcaaaatccct 420
Db 55116 AATCTGATCCAAATCCGGTGTGTGACATGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 55057
QY 421 gccgtcgtatccgagccgcttcgcacatctcgtcaatgcgatgacacaaatccgttngcg 480
Db 55056 GCGTGCATGCGCGAGCGCGTTCGCGCATCTTGTCAATGCGATGAGACCAATTCGCTGGCG 54997
QY 481 gcgacccgtgtgttgcgtacaaagacccgngangattcagaagangtngctgtgta 540
Db 54996 GCAAGCCCTGTGTGTGATCAAAAGAGCCGCGAGATTTCAACAGAGTTTGTCTGTGTA 54937
QY 541 ttgagccgtttgacgcgagcgttaaatccatgctgtgtgaagcagctgcgagacgttgcg 600
Db 54936 TTGAGCCGCTTGAACCGAGCGTAAATCCATGTGTGAAGCAGGTGGCGAGACGTGCCG 54877
QY 601 tctgaaatctctgcacacatcgaacacatgaaatcggcgccgactccgcccgtgttg 660
Db 54876 TCTGAAATCTCTGCACATCGAACAACATGAATTCGGCGCCGCGATCCGCGGTTTG 54817
QY 661 agtggcagcacatcattcatctgaagcgcgtgcgtgcaacaaacgcttttgacatc 720
Db 54816 AGTGCAGCAGCATCATTTTCATTGACCGCGTGCCTGCAAAACAAACCGTTTGGACCATC 54757
QY 721 aatatacaagatgaattgcacatcgcagcttgcgttgaacagcgctctgaacccgag 780
Db 54756 AATATCAAGATGAATTTGCGATCGAGCGTGTGTTTCAACAGCGCTCTGAACACCGAG 54697
QY 781 cggctgattcgttgggtgttctcaagtaacaacaaacacgctctgtgtaaccgtttg 840
Db 54696 CGCGTATTCGTTTGGGTGTTCTCAAGTCAACAAACACGCGCTTGCCTACCGTTTGG 54637
QY 841 ggtgcgaagatcgcgaatctcgcgagcgaattggttgacgagacaaacgcgctgatt 900
Db 54636 GGTGCAGAAATGCGCAAAATTTACTGCGCGGAATTTGTTGACGAGCAACCGCGTGATT 54577
QY 901 tccggttcggtatgtaacgcgcgcatcaacaagcgcgacagatatttggagcgatc 960
Db 54576 TCCGTTTGGTATTTGAACGCGCGCATTAACAAGCCGCGACGATTTATTTGGGCGCTAC 54517
QY 961 cacatcaagattccgctatcgaagaagcgcgcgacgaagagcgttcgctggtgtgctg 1020
Db 54516 CACAAATGAGTTTCCGTTATTCGAAGAAGGCGCGCAAAAGCGCTGCGCTGGGTTGCG 54457
QY 1021 ccgagccgagacaaatactcaatcagcgtaagcgttcgcttccgttgaagcaaaacaa 1080
Db 54456 CCGAGCGGAGCAAAATCTCATTCACGCGTACGACGACCTCGGCCATTTTCGAAAAACAAA 54397
QY 1081 ctctcaagttcagacagcgttcaacgctgagcgagcgccatggtgacgttgcattgact 1140
Db 54396 CTCTCAAGTTCAAGACAGCGCGTCAACGAGTGGCGACGCGCCGCAATGATGATGACT 54337
QY 1141 tacgagcgttaatgcccgtagaaatcccttaacctgttgcggaattaatgctc 1200
|||||


```
Db 54336 TACGAGCGCGTAATGCGCGTAGACATCTGCTACCGCTGTTTGGCGATTATATGTC 54277
Qy 1201 ggcagatccgacagcgcgcaagcattgggtctgttggaattggaagaagacccctcgct 1260
Db 54276 GGCAGTACCGACAGCGCCGCAACGATTTGGGTTGCTTGGATTGGACGAGAACCTCGCT 54217
Qy 1261 ttgtgcaacttcgtctgtccgagcaataacgaatgagccgcgtgtgttgtaagtgctg 1320
Db 54216 TTGTGCACTTGTCTGTGCGCGGCAATAATGAGATACGGCCGCTGTGTGCTTAAGTGCTG 54157
Qy 1321 gaaacnttgagaagaagcttga 1344
Db 54156 GAACCATTTGAGAGAGAGCTGA 54133

RESULT 3
LOCUS A96092 1344 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 125 from Patent WO924578.
ACCESSION A96092
VERSION A96092.1 GI:6779938
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1344)
AUTHORS Pizsa, M., Scarlato, V., Rappuoli, R., Grandi, G., and Masignani, V.
TITLE Neisseria meningitidis
PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);
JOURNAL PIZZA MARIAGRAZIA (IT); GRANDI GUIDO (IT); MASIGNANI VERGA (IT);
FEATURES
    location/Qualifiers
        1..1344
            /organism="unidentified"
            /db_xref="taxon:32644"
BASE COUNT 329 a 366 c 365 g 284 t
ORIGIN

Query Match 91.5%; Score 1229.4; DB 6; Length 1344;
Best Local Similarity 94.2%; Pred. No. 5.8e-289;
Matches 1266; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

Qy 1 atgatttaataataaaaggtctaaactcgtccatcgccgagcgagaccggaagcaatt 60
Db 1 ATGATTAAATCAAAAAGGTTCTAAACCTGCTCCATCGCGGCGAGACCGGAGCAAGCCGTT 60
Qy 61 tatgacggcgccgtcatctaccgaagtcgctgtctgtgacgaagaatactgcggtatgcg 120
Db 61 TACGAGCGCGCCGCCATTACCGAAGTCGGTGTGGCGAAGAAATAGCCGATATGCGC 120
Qy 121 cccttgaatgaagaatcaagaagcgatgcgtcaaaaaaagccaagtgcgttgaagac 180
Db 121 CCTTCATGAAGTCAAGGAGGCGATGCGCTCAAAAAAGGCCAAGTCTGTTGAAGAC 180
Qy 181 aaaaagaatccggcggtgtgtttacccgcncttcaggcaaaatccgcccacatcat 240
Db 181 AAAAAAATCCGGCGGTGTGTCTTACTGCGCCGCTTCAGGCAAAATCCGCCGCAATTAC 240
Qy 241 ccgagcgaaaagcgctacttcactcagtcgctgtgattgcgtltgaaggaacgaatc 300
Db 241 CGTGGGAAAGCGGCTACTTCAGTCAGTGCTGATTCGCCGTTGAAGCAACGAGCAATTC 300
Qy 301 gagtgcgaacgctacgcgcgcgaagcgttgcaacttaagcggcganganatmgmc 360
Db 301 GAGTTTGAACGCTACGACACTGAAGCGCTGCAAACTTAAGCGGCGAAGAGCTGCGCGC 360
Qy 361 aatcgtaccaatccggttgtgactgcgtgcgtanccgctcgtltcaagaataaccct 420
Db 361 AACCTGATCAATCCGGTTGTGTGACGCGCTGCGCAOCCTCGCTTACGAAATATTCCT 420
Qy 421 gccgtgcgtgcgagccgctgcgcacatctcgtaacgcatgtaagaaccaatccgctng 480
Db 421 GCCGTGATGCCGAGCCGCTTGCCCATCTTCGTAATGCGATGAGACCAATCCGCTGACT 480
Qy 481 gcaagacctgtgtgtgtatcaagaacgcnngatattcaagcgangtntgctgta 540
Db 481 GCCAGCCTTAGGTCATTATCAAAAGACGCCGAGGATTTCAAACGCCGCTGTTGGTA 540
Qy 541 ttgagccgttgcagcgagctaaatccatgctgttaagcgagctgycgcaagctgcg 600
Db 541 TTGAGCCGTTTGACCGAAGCGCAAAATCCATGTTTGAAGGACGCTGGCGACAGCGCGC 600
Qy 601 tctgaanaatgctgcacacatgcgaacacatgaaatccgagccgcgaatccgcggtttg 660
Db 601 TCTGAANAATGCTGCACACATCGAAACACATGATTCGGCGCCGCTGCTCGGCTTTGG 660
Qy 661 agtgcagcgacattcattatctatgagccggtcgtgtgcaaaaaacgcttgagaccatc 720
Db 661 AGTGCGACGACATTCATTCATTCGACGCGGTGCGGCGCAATAAACCCTGTGGACATTC 720
Qy 721 aattcaagatgtaattgcacatcgagcgttgtltgcacagcgccgtctgaacaccgag 780
Db 721 AATTTCAGATGTAATTACATTCGCGCTTGTGTCGAACAGCGCGTCTGAACACCGAG 780
Qy 781 cgcgtgattgcttgggtgtgtctcaagtcacaacaacacagccctctgcgtacgctttg 840
Db 781 CGCGTATTGCCCTAGTGTGTTCTCAAGTCAACAAACCGCGCTCTTGGGTACCGTTTGG 840
Qy 841 ggtgcgaagaatgcgaatctactgcggcggaattgtgtgaagcgagacaacgcgctgatt 900
Db 841 GGTGCGAAAGTATCCCAATTCCTGCGGCGAATGTGTGTGACCAACACCCCGTGATT 900
Qy 901 tccggtcgtgattgaacgcgcgatcaacaagcgcgcgacgcatatttggagacgtac 960
Db 901 TCCGTTGGGTATTAACGCGCGCGATTAACAAAGCGCGCACGATATTGGAGACCTAC 960
Qy 961 cacaatcgaattccgctatcgaaagaagcgcgagaaagagctgtcgtcgtgtgtgcg 1020
Db 961 CACAATCAGATTTCGTTATCGAAGAGCGCGACGCAAGAGCTGTGCGCTGCGTTGCG 1020
Qy 1021 ccgagcgcggaacaaatcctacatcgcgtagcagccctcgccattccctgaaaaaaa 1080
Db 1021 CCGAGCGCGAATAATCTCCATCAGCGCTACAAACCTCGCCATTTCGAAAAACAA 1080
Qy 1081 ctcttcaagtlcaacgaacgcgtlcaacggtlbgacgcgcgtatgtccgattgttact 1140
Db 1081 CTTTCAAGTTCAACACACGCGTCAACGCGGCGACCGCCATGTGCGCATTTGGTACT 1140
Qy 1141 taagagcgctaatcgctgtagaataccctgacacctgcttttgcgagatttaactgtc 1200
Db 1141 TACGAGCGGCTGATGCCCTTGATATCTGCCCCACCTGCTTTGCGCGATTATATGTC 1200
Qy 1201 ggcagatccgacagcgcgcaagcattgggttgcatttgaaattgacgaagaacctcgt 1260
Db 1201 GCGGATACCGACAGCGCGAGCGATTTGGTGTCTTGAATTGAGCAAGAACCTGCT 1260
Qy 1261 ttgtgcagcttcgtctgcgcggaataacgaatlangcccgctgtgtgcgtaagtgctg 1320
Db 1261 TTGTGAGCTTGTCTGCGCGGCAAAATACGATACGAGCCGCTGTGTGCGCAAGTGTCTG 1320
Qy 1321 gaaacnttgagaagaagcttga 1344
Db 1321 GAACCATTTGAGAGAGAGCTGA 1344

RESULT 4
LOCUS AE002412/c 12273 bp DNA linear BCT 25-MAY-2000
DEFINITION Neisseria meningitidis serogroup B strain MC58 section 54 of 206 of
ACCESSION AE002412 AE002098
VERSION AE002412.1 GI:7225783
KEYWORDS
SOURCE Neisseria meningitidis MC58.
ORGANISM Neisseria meningitidis
```


LOCUS AX044029 349980 bp DNA Linear PAT 24-NOV-2000
DEFINITION Sequence 108 from Patent WO0066791.
ACCESSION AX044029
VERSION AX044029.1 GI:11342913
KEYWORDS
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis.
REFERENCE 1 (bases 1 to 349980)
AUTHORS Pizzo, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J. C.,
Masignani, V., Galeotti, C., Mora, M., Ratti, G., Scarlato, V.,
Scarlato, V., Rappuoli, R., Frizer, C. M. and Grandi, G.
TITLE Neisseria genomic sequences and methods of their use
JOURNAL Patent: WO 0066791-A 108 09-NOV-2000;
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
Source location/Qualifiers
1..349980
/organism="Neisseria meningitidis"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces -seq 1: 1 to
349980 349980 bases -seq 108: 300001 to 649980 349980
bases -seq 109: 600001 to 949980 349980 bases -seq 110:
900001 to 1249980 349980 bases -seq 111: 1200001 to 1549980
349980 bases -seq 112: 1500001 to 1849980 349980 bases -seq
113: 1800001 to 2149980 349980 bases -seq 114: 2100001 to
2272325 172325 bases"
BASE COUNT 84410 a 84863 c 94187 g 86520 t
ORIGIN
Query Match 91.5%; Score 1229.4; DB 6; Length 349980;
Best Local Similarity 94.2%; Pred. No. 6,3e-289;
Matches 1266; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 1 atgatttaataataaaaaagtttaaaccttccatcgcgagcagccgagcaatgacatt 60
Db 297909 ATGATTAAATCAAAAGAGTCTAAACCTGCCATCGCGGAGACCGGAGCAAGCGTT 297850
QY 61 tatgagcgcccgctcatctaccgaagtcgctgcttgaggaaataatgctgcatcgc 120
Db 297849 TACGAGCGCCCGGCCCTATTACCGAAGTCGCGCTTCTTGCGCAAGAAATATGCGGATATCGC 297790
QY 121 cccttgaatgaagaatgaagaagcgatgctgcgtcaaaaaaagcgcaagtgcgtgtgaagac 180
Db 297789 CCCTGATGAAGGTCAAGGAGGCGATGCGCTCAAAAAAGCGCAAGTGTGTTGAAGAC 297730
QY 181 aaaaagatccggcgctggtgttttaccgcgccngtttcaggcaaaatcgccgcacatccat 240
Db 297729 AAAAAAGATCCCGGGCTGTGTTTACTTGCGCGCGCTTCAGGCAAAATCGCGGATTTCAC 297670
QY 241 cggcgagaaagcgctactcactcagtcgctgctgctgtgaagcgcaagcgaatc 300
Db 297669 CGTGCGCAAAAGCGGCTACTTCAAGTCAGTGTGATTTGCCGTTGAAGCAACAGCAAGCAATC 297610
QY 301 gaattcgaacgttaacgcccgaagcgctgtgcaaaccttaagcgcgagaaantnnngnc 360
Db 297609 GAGTTGAACGCTACGACCTTAAGAGCGTGGCAAACTTAAGCGCGCAAGAAAGCGCGCGC 297550
QY 361 aatcgtacccaatcgctgttgcgactgcgtgcgtgancgctgcgttcgaacaaatccct 420
Db 297549 AACCTGATCCATCCGGTTTGGGACTGGCTGCCAACCCTCGCTTCACCAAAATTCCT 297490
QY 421 gccgtcgatgcgagcgcttgcacatcttcgtcaatgcatggaacccaatccgttncg 480
Db 297489 GCCGTGATGCGAGCGACCGCTTCCCATCTTCGTCATGCGATGAGACCAATCCGGTGGCT 297430
QY 481 gcagacccgtggtgtgatacaaaagcgctgacgagatcctcaagcagantgtgtgta 540
Db 297429 GCCGACCTTACGGTATTATTAAGAAAGCGCGCGAGGATTTCAAGCGCGCGCTGTGTGA 297370
QY 541 ttgagccgttgagccgagcgttaaatccatgctgttgaagcagcgtcgacagcgctgacg 600
|||||

Db 297369 TTGAGCCGTTTGAACCGAGCAAAATCCATGTTTGAAGCGACGTGGCGAGCGTGGCG 297310
QY 601 tctgaaaatgctgccaacatcgaacacatgaatttcgagcccgccgacatccgagctgttg 660
Db 297309 TCTGAAAATGCTGCCAATCGAATCGAAACATGATGTCGGCGGCCGATTCCTCGCGTTTG 297250
QY 661 agtgcagcgaatcattcattcattgagccgctgcgtgtaacaaacacgcttggacatc 720
Db 297249 AGTGGCAGCGCACATTCATTTCAATCGAGCCGCTGGCGCGCAATAAACCCTGTGACATC 297190
QY 721 aattacgaatgtaattccatccgacgacgttttttgcacaagcgcttcgaacacgag 780
Db 297189 AATTATCAAGATGTAATTAATACATTTGGCCGTTTGTTCGAACAGCGCTGTGAACCGAG 297130
QY 781 cgcgtgattgcttgggtgtgttcacagtcgaagtcgaacaaacacgacgtctgtgtacgcttg 840
Db 297129 CCGGTATGCGCTAGTGTGTTCTCAAGTCACAAACCGCGCTCTTGCGTACGTTTG 297070
QY 841 ggtgcgaagatcgcgaatattctgctgagcggaattgtgtgagcagacaacgcgtgatt 900
Db 297069 GGTGCAAAAGTATCGCAAAATTAATCTGCGGCGAATTTGTTGACACAGCAACCGCGTGTATT 297010
QY 901 tccggttcgtaattgaacgagcgcgatcacaagaagcgcgacgattatttggagcgtac 960
Db 297009 TCCGCTTCGCTATTGAACGCGCGATTAACAAAGCGCGCACATTTATTGGAGCGCTAC 296950
QY 961 caacaatcgaatttcgtaattcgaagaagcgcgagcaaaagcgtgttcgctgtgtgtcg 1020
Db 296949 CACAATCAGATTTCCGTTATCGAAGAGCCGCGACGAAGAGCTGTTCGCTGGTTGCG 296890
QY 1021 ccgagcgcggaataatctccatccatcgctgacgaccttgcgacatttctcgtaaaaaaca 1080
Db 296889 CCGAGCGCGAATAATCTCATCTACGCGTACAACTTCGCGCTGCTGAAACAA 296830
QY 1081 ctcttcaagttcacagcagcgctcaacggttgcgacgagcgacatgttgcgattgtgtact 1140
Db 296829 CTCTTCAAGTTCAACAGCCGCTCAACGCGCGGAGACCGGCATGTGCGGATTTGGTACT 296770
QY 1141 taagcgcgctaatgctgcgttagacaatcctgctaccctgtcttgcgagatttaactcgt 1200
Db 296769 TACGAGCGCGTATCCCTTGATATCCGCGCACCCCTCTTTTGGCGCATTTAAATCGTC 296710
QY 1201 ggcgaatccgaacgagcgcaaacatltggtgttcgttgaaattgagcaaaacacccgct 1260
Db 296709 GCGCATACCGAGCGCGCCAGCATTTGGGTCCTTGGAATTTGACCAAGACCTCGCT 296650
QY 1261 ttgtgcagcttcgttcgcccgggcaatacgaatannngccgctgtgtgtaagtgtcgt 1320
Db 296649 TTGTGCACCTTCGTCGCGCGCAAAATACGAATACGCGCCGCTGTGCGCAAAAGTGTG 296590
QY 1321 gaaacnttggagaaagagcgtga 1344
Db 296589 GAAACCATGGAAGAGAGGCTGA 296566
RESULT 6
A96098
LOCUS A96098 1344 bp DNA Linear PAT 07-SEP-2000
DEFINITION Sequence 131 from Patent WO9924578.
ACCESSION A96098
VERSION A96098.1 GI:6779941
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1344)
AUTHORS Pizzo, M., Scarlato, V., Rappuoli, R., Grandi, G. and Masignani, V.
TITLE Neisserial antigens
JOURNAL Patent: WO 9924578-A 131 20-MAY-1999;
PIZZA MARIAGRAZIA (IT); SCARLATO VINCENTO (IT); RAPPUOLI RINO (IT);
CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)
FEATURES
Source location/Qualifiers
1..1344

VREGDVKKQVLFEDKKNPGVFTAPASGTAVTINRGEKRVLOSIVIVEGDEQITF
 TFEAOLASLAEQVKNLIESGLTARTRFESVPLDALPSSIPFNAMDTBLA
 APEVLYKEFEDKDLVLRLEFNGCKPVLKPADSNIPSPLEGTIKSFSGV
 HPAGLVGTHIEFVDPGATKQVWHLNIOVIAIKFTTGELTETDILSLAGOVKNP
 RIVRRFLGNLSQITANENLAGEBNRIVISGNAAPVPIGYSIAOVSIAGR
 EEELHGMIPGSDKRSITRTVYGHBEKRIFFNTYVHGEBRANVPYIGAEIRVPLDI
 PHLRLDLAAGTDSQNLGCLDEDEDLALCTYVPGKNNKPMRLADEKLEKG*
 1462..1499
 terminator 432 a 299 c 375 g 452 t
 BASE COUNT 432 a 299 c 375 g 452 t
 ORIGIN

Query Match 41.8%; Score 562; DB 1; Length 1558;
 Best Local Similarity 64.8%; Pred. No. 2.7e-126;
 Matches 874; Conservative 0; Mismatches 467; Indels 12; Gaps 3;

1 atgattaaatacaaaaggctcaaacctgcacatcgagcgagacggagcaagtcatt 60
 95 atgattaaatacaaaaggctcaaacctgcacatcgagcgagacggagcaagtcatt 154
 61 tatgacggcgccgctcattaccgaagtcgctgttgagcaaaataatgcggtatg 120
 155 catgacggcgccgctcattaccgaagtcgctgttgagcaaaataatgcggtatg 214
 121 cccatgataaagtcgaaggaagcgatgcgctcaaaaagcgcaagtcgctgttgag 180
 215 ccttcaatgaatgctgtgtaagcgatgtaaaagcgcaagtcgctgttgaggaat 274
 181 aaaaagatccggcgctgtgttaccgagcngttcagaacaaatgcgcgcattcat 240
 275 aaaaagatccggcgctgtgttaccgagcngttcagaacaaatgcgcgcattcat 334
 241 cgcgcgaagaagcgctgactcagtcgctgttaccgagcngttcagaacaaatgc 300
 335 cgcgcgaagaagcgctgactcagtcgctgttaccgagcngttcagaacaaatgc 394
 301 gagtcgaacgctacgagcgccgaagcgctggaacaaatgcgaagcgagaaatnngnc 360
 395 acccttactgcatgaaagtcgcaatgagcgtgcacatgcacgagcaagaaacaa 454
 361 aactcgaacatccgctgttgagtcgctgttaccgagcngttcagaacaaatgc 420
 455 aactcgaacatccgctgttgagtcgctgttaccgagcngttcagaacaaatgc 514
 421 ggcgtcgaatcgagcgctgttcacatcttcgtaatgagtcgaacaaatgcgctg 480
 515 gcatgtgacacgcatgctcgtatcattcttgcataatgacaaatgcgctgagcg 574
 481 gcaacccctgtgttgatcaagaagcgagcngatccgaagcngatnngtctgta 540
 575 ggcgagtcggaatgctgtttaaagaaatgcgaacaaatgcgctgagcgttactg 634
 541 ttagacgcgtttagcagcgagctaaatgcgaacaaatgcgctgagcgttactg 594
 635 ttagacgcgtttagcagcgagctaaatgcgaacaaatgcgctgagcgttactg 694
 595 gtcgcgtttagcagcgagctaaatgcgaacaaatgcgctgagcgttactg 651
 695 atcccgctttagcagcgagctaaatgcgaacaaatgcgctgagcgttactg 754
 652 ggcgcgtttagcagcgagctaaatgcgaacaaatgcgctgagcgttactg 711
 755 ggcgcgtttagcagcgagctaaatgcgaacaaatgcgctgagcgttactg 814
 712 ttagacgcgtttagcagcgagctaaatgcgaacaaatgcgctgagcgttactg 771
 815 ttagacgcgtttagcagcgagctaaatgcgaacaaatgcgctgagcgttactg 874
 772 aacacgaagcgctgagctgtgttgagtcgaacaaatgcgctgagcgttactg 831
 875 ttagacgcgtttagcagcgagctaaatgcgaacaaatgcgctgagcgttactg 934
 832 accgctttagcagcgagctaaatgcgaacaaatgcgctgagcgttactg 891

Db 935 ACAGGCTCTTGGTGGGAATCTTCCCAATTAACCGCAATAGTAAATATGCCGGTGAAC 994
 892 CGGCTGATCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 951
 995 CGGCTGATCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1054
 952 GAGCGCTACCAATCAGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT 1011
 1055 GCGCGTACCAATCAGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGAT 1114
 1012 TGGCTGATCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1071
 1115 TGGCTGATCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1174
 1072 AAAAAAATCTCTCAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 1131
 1175 AAAAATCTCTCAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 1231
 1132 ATGCTGATCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1191
 1232 ATGCTGATCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1291
 1192 TTAATCTCTCAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 1251
 1292 TTAATCTCTCAATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGAT 1351
 1252 GAGCTGATCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1311
 1352 GAGCTGATCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1411
 1312 AAGCTGATCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1371
 1412 GAGCTGATCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1471

RESULT 8
 AE006171 10563 bp DNA linear BCT 08-MAR-2001
 LOCUS Pasteurella multocida pm70 section 138 of 204 of the complete genome.
 DEFINITION Pasteurella multocida pm70 section 138 of 204 of the complete genome.
 ACCESSION AE006171 AE004439
 VERSION AE006171.1 GI:12721691
 KEYWORDS
 SOURCE Pasteurella multocida.
 ORGANISM Pasteurella multocida.
 Pasterella multocida.
 Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Pasteurella.
 REFERENCE 1 (bases 1 to 10563)
 AUTHORS May, B.J., Zhang, Q., Li, L.L., Paustian, M.L., Whittam, T.S. and Kapur, V.
 TITLE Complete genomic sequence of Pasteurella multocida, pm70
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)
 MEDLINE 21145866
 REFERENCE 2 (bases 1 to 10563)
 AUTHORS Zhang, O. and Kapur, V.
 TITLE Direct Submission
 JOURNAL Submitted (24-OCT-2000) Department of Veterinary Pathobiology, University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN 55108, USA

FEATURES
 source
 1..10563
 location/Qualifiers
 /organism="Pasteurella multocida"
 /strain="pm70"
 /db_xref="taxon:747"
 94..1605
 /gene="PM1326"
 94..1605
 /gene="PM1326"
 /codon_start=1
 /transl_table=11
 /product="unknown"
 /protein_id="AAK03410.1"

gene
/db_xref="GI:12721692"
/translation="MMIDVMTQRDCILKLKNICKSEFDPVPLKNDVLDIYAGEHCLV
GENAGSTLICKIATPIPRPGQIEPKESEKSEPTVKEAOSIGIGITHOELMIVOL
SVLENIFLGNKEITVFCGKMDQTMHKKTOIISELELDDPPAKVADLSIAQOOMEI
AKAVESEYKVIIFDEPTSSISKRNTOTLEKIIHQLEKKNAMITISHRLDEFYIADQ
VTVLRDGIKRTGMAYRDTSPDEIVKLWGREINFAEYHNYLPTREVLVESTJTNKH
KSSSEFLKEGELGAGVAGTELLRAIFGADENSGAIFYLNAQVEIRSPDAAVO
KITGFTEDRKOOGELIGLSIRENITLPLIKRPNWGLNKOERLVAEENROKRLIYS
HSQPOISOTLSLPGNOOKVIARWLESQVYLIPEDEPTRGIDGAKSEIDIMQFIDN
GGATVWSSDLPELITMSDRVIVMROQGIYREVKKAEMTEBNMLRMIGVDE"
1598. .2641
/gene="tbsc_3"
/note="PM1327"
1598. .2641
/gene="tbsc_3"
/codon_start=1
/transl_table=11
/product="tbsc"
/protein_id="AAK03411.1"
/db_xref="GI:12721693"
/translation="NNRCMTNLKRIKIFSLGIGITLLMIGMSLSDVFLSTNNIIN
ILQVSIICVISGMYVYILTGIDLSVGSIVALSAVCLGFTHMGMDLGENPSEFS
VLMIYISAIATIFVIGILGVYNGLVIVGKYTPITTLGMMGIARGLATITSDGKI
YNFPTLRFILGNGRIALTETFPVPIIALIVLIVLVSFYLTQTMFGROIYALGNRE
AVRLSGINIEKLEIKAVYINGALAAIGAILIVGRLNAOPIDAGNGELDAIVIGG
TSLMGVGSVSTSGALIMGYLQNGLTLLNTSYLQRLITIGLVIILAVLDQLRGE
VSTRKLKRFEPK"
3187. .4527
/gene="PM1328"
3187. .4527
/gene="PM1328"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK03412.1"
/db_xref="GI:12721694"
/translation="MITTKKGLNLPISSGSEQVIRDGNAITTEVALLGEYVGNRPSNK
KRESDVKKGOVLFEDKNPNPVFTAPASGTVAIHRGAKRVLOSVIKIEGNOQITF
EYETREOLNOLTSBOYRONLOASGLMTALTREPSKVPADATPVSIENVAMDNPIPC
ADPOYIVQOSAOAFEGCEIVLSRLHGGKYVLCANANASIPSPSIANDYERPNRPHN
GLSTGHIHIDPVSAKFEWYINYOVIYVAGLFTTGLEDVSRVSLAGVQVKRFLV
RVLIGANLSQLTANEGKDEBNRIVSGSVLSGAKAAGPVYDLYGRALVOSVLEGRKE
FLGIMPGANKRYSLSRTVIGHFSKLFNFTTALNGERAMVPIGAVERVMDLIIPPL
LLRLAAGCTDGAQAALGCLLEDEDLALCTFVCPGKNEYGPLRLQALDKTEKES"
4530. .5762
/gene="nqrb"
/note="PM1329"
4530. .5762
/gene="nqrb"
/codon_start=1
/transl_table=11
/product="nqrb"
/protein_id="AAK03413.1"
/db_xref="GI:12721695"
/translation="MGIKHFIEKIEPAFLPGKYEKNVALFEATATFLYTPGVYTHKA
SHVRALDSKRMVVLWLSLEPFAMFEGMYVGAOAILATDALGTLQOALISGHMOYALS
HILGADITLSAGASKMILGATYELPLIYLIPLVGGFWELFPAVRKHHEINSEFVVS
ITLALIVPTLPLMQAALATTPGCVVAKETFGVCGKRNPNPALAARAPLEPAPAOIS
GUTVWVADGFSGATALSOWAVGEGELKHVVTGOPTIMDAITGNIPTSGIEYSTLA
LIGAIIVFARIASWRILAGVIMGIMATSNLNLIGSSINPLFAMPWYWHFVIGFA
LGMFWATDPVSAAFPNKGMWYGIILGAMAVLIRVNPAYPEGMMLAIIIFANLFAPY
FDYLVQANIKRRRARNV"
5755. .6537
/gene="PM1330"
5755. .6537
/gene="PM1330"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK03414.1"
/db_xref="GI:12721696"
/translation="MFRNKDSVCGTLLVITILLISAGSIYAGSAVLLKPTOIEQELD
KKNILISVAGLLQPTTKNSRIKTIYANNIEARLVDLNTGDFAPAPGPDAAKAVKNPA

gene
/db_xref="GI:12721697"
/translation="MADRKLLGILLSPYMDNPIALQILIGCSALAAVTKLETALIV
TFAYIVFAEISNLFISLIRNYIPNSVRIIVQMAITASIVILVDQILIRAYAGLSQLS
VEFGIITNCVLMGRSAEAPAKMSQPLESEFDINGLGYGALIVSAFIRELIGSGKL
FGMTVFOTIDGGMWQTNGLFLAPSAFFIIGIIMGIFRTLPNQVEK"
7160. .7756
/gene="nqre"
/note="PM1332"
7160. .7756
/gene="nqre"
/codon_start=1
/transl_table=11
/product="nqre"
/protein_id="AAK03416.1"
/db_xref="GI:12721698"
/translation="MEHYISLPKSVFIEFNMLSPFLGMCFTLANSKRYSTAFGLGIA
VIVVIGIAPVNOIYVSLFIKDSALVOGIDLSFLNFTFYGIVIAIVOLLEMYLDYF
PALYNALGIFPLPLTVNCAIREGVSFPMQRODTPVESYVYIGAGIGMMLAIVALLAGI
TEKMYADVPAGLRGIGITFVGLMALGFSFSGIQL"
7772. .8995
/gene="nqre"
/note="PM1333"
7772. .8995
/gene="nqre"
/codon_start=1
/transl_table=11
/product="nqre"
/protein_id="AAK03417.1"
/db_xref="GI:12721699"
/translation="MEITGIMFPIYVIALAVIILFAKSKLVNSGDITIEINDPSK
AIIHPAGGLGALASOGIFVVSACGGGSCCOCITYVEGGDILPTLSHISKREPA
KEGYRLSCOVNPNKMKVELPEEVEGVKWECTVLSNDKATITIKELKALIPGEYEP
FRAGGYIOIEAEPTVAAYKDFDIPBEYHBDWKYNLMRYVSKYDEHIIRAYSNASTPE
EKGIIMLVNRIATPPPNPDADPGQMSYIWSLAKADKVTISGPFGEFKAETDAEYV
FYVGAGAGAPMRSHTIEDOLRLKRSKMSFWGARGSKREMFYEDPTLOAENDNFW
HVALSDPQGDMDGCTGTGITHVLYENYIKOHAPBEDCEYKMGPIINNASYIKMLKD
LGVEDENILDDFG"
9164. .10204
/gene="apbe"
/note="PM1334"
9164. .10204
/gene="apbe"
/codon_start=1
/transl_table=11
/product="apbe"
/protein_id="AAK03418.1"
/db_xref="GI:12721700"
/translation="MKHIINSIIIGCSILACHKSPEVYVITGKTMGTYSYKVID
DSQIKIPTLEHNGQIEIIEKYNKRSYTLSTSELSQFNQNTQIDNPVEISADLAPVL
KEAIRLHGVTBEGALDVTPIVNLWGFGEKROPNTEBOVKKRAMVTEKKKLTIE
OGDKFYIAKTVPOLYIDLSIAKGFQDVVAHLAVKGLANYIVETIGETINAKGNIE
GKDMOIAIEKPNFEDGSRASQIIGLDFSMATSGNRYNFEENGRRFSHIDIKRTGTP
IEHRLASTIVLADSTWTADQLSTGLVYLGEKALEVAEKYNLLVYLISKTEKGEFAAM
SSAFKELINEOH"
10214. .10468
/gene="PM1335"
10214. .10468
/gene="PM1335"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK03419.1"

REMARK The H. Influenzae sequence has been updated by R. Fleischmann. New database matches have been assigned, product names have been improved, and a number of frame shifts have been corrected. We gratefully acknowledge the work of Talusov et. al. We have incorporated their annotation into the /notes fields of the corresponding H. influenzae genes

REFERENCE 5 (bases 1 to 13299)

AUTHORS White, O., Clayton, R.A., Kerlavage, A.R., Fleischmann, R.D., Peterson, J., Hickey, E., Dodson, R. and Gwin, M.

TITLE Direct Submission

JOURNAL Submitted (28-May-1998) The Institute for Genomic Research, 9112 Medical Center Dr., Rockville, MD 20850, USA

REMARK The whole genome was shifted by 588 nucleotides for a new start On Sep 30, 1996 this sequence version replaced gi:1220812.

FEATURES

source Location/Qualifiers

1..13299

/organism="Haemophilus influenzae Rd"

/db_xref="taxon:71421"

/complement(85..1455)

/gene="HI0161"

/complement(85..1455)

/note="similar to SP:P06715 GB:M13141 PID:146248

PID:466637 GB:000096 percent identity: 74.17; identified by sequence similarity; putative"

/codon_start=1

/transl_table=11

/product="glutathione reductase (gor)"

/protein_id="AAC21833.1"

/db_xref="GI:1573119"

/translation="MTKHVDYIAIGGSGGSIASINRAASYGKCAIIEAKHIGSTCVN VCCPKVMVEYGAHIAEINNVAPDGFDEVKKDFSKLIESROAISRHTSENNY LAKNNIDVINGEKKFVDADHTIEVLTADGTEQYADHIIATGSRPAPRNKGGEEYI DSDPAALTELPERRAIVGAGYIAVELSGVLSNCSYETHLILVRHAPMRNDPIIVET LVEVLAQDGIQLHTNSTPSEIVKNAADSLVRCDSGDVTVDCVIMAGRPVTDKIG LKNAVETNEHGIVKDYQNTNVKGIYAVAGDIIENGIELTPVAVAGRPVLSERLFFNR KPLEYLDVSLVPTVSHPIGTGVLTEPQAIQGYGAENKVKVKSFTAMTAVTQHR OFCKMKVCVKDEKVKVGLHGIGFVDEMIGQFAVAIKMGATKADFNTVAIHPITGSE EFTVNR"

/complement(1671..2270)

/gene="HI0162"

/complement(1671..2270)

/note="similar to GB:000096 PID:1773118 PID:1786637

percent identity: 25.82; identified by sequence similarity; putative"

/codon_start=1

/transl_table=11

/product="lipoprotein, putative"

/protein_id="AAC21834.1"

/db_xref="GI:1573120"

/translation="MALVNRKIKTLSSVGIILATFLAGCOAOSNIILAFPPASASMN VNRRAVSTTRKSRALIOELASTYTKHRELKLNKSPSYTOLFQOVMOONLISKFRVQ ULNGSNAVTVDFREFGTOVEGSLNRKLNKTKIOATYVVOGAKGSYNKSFNVTHSQEG VFNAGNDEIHKVLSQTFNDIYNNIYDQEVAAALINQYSN"

2364..2675

/gene="HI0163"

2364..2675

/note="HI0163"

/note="similar to PID:663269 percent identity: 46.53; identified by sequence similarity; putative"

/codon_start=1

/transl_table=11

/product="transcriptional regulator, putative (bolA)"

/protein_id="AAC21835.1"

/db_xref="GI:1573121"

/translation="MSTIQILIEOKIQEPQPHLAIENESHLHNSRGSESHFKCVIY SADFKNIRKQVQRHRIYQLINBELNHSIHALLALFTPEEKAKQNETVPSTKCAICIG R"

2966..4309

/gene="HI0164"

2966..4309

/gene="HI0164"

/gene="HI0164"

/note="similar to PID:663269 PID:663270 percent identity: 64.96; identified by sequence similarity; putative"

/codon_start=1

/transl_table=11

/product="NADH:ubiquinone oxidoreductase subunit A, Na translocating"

/protein_id="AAC21836.1"

/db_xref="GI:1573122"

/translation="MTTIKGLDLPIAGKPAQVYTHSGNVAIVNOVAIIIGEYVGRPSMK VREGVAVKGOVLEFDKRPQVITRAPASCTITAIRGKERVQSVYVINEGEKLTIE AKYSTEQLWTLTSEEDVKONLISGLMTALTRPFSKVPSTIESASSIFVNAAMTNPILA ADPSVLAKEYSDQFTNGLVLSRLPFSKPLHLCKADSNIPADLENLIDHTGVHP AGLVGTHHIFIDPVGIOKTVWHINXODYIAGVGLFTTGGELSERVLSLQAPQKEPRL VRTTIGANISOLTONELISAGKNRIVISGVLCQIAKSDHLYDRVALOVSVIAGNEK BEFGWIMPOANKSVTRYTVLHGHSKRLNFTTSENSEGERMAVPIGSEYERVPDILPT LILRLIDVTPGAGDELGLDEBDLALCSFVCPKRYEGSLRQVLDKIEREG"

4312..5547

/gene="HI0166"

4312..5547

/gene="HI0166"

/note="similar to PID:663269 PID:663271 percent identity: 71.33; identified by sequence similarity; putative"

/codon_start=1

/transl_table=11

/product="NADH:ubiquinone oxidoreductase, subunit B (nqrB)"

/protein_id="AAC21837.1"

/db_xref="GI:1573123"

/translation="KGLKNLFEKMEPAPLPGKYSKLVPIFESITLLYTPRTVTHKN THVRALDSKRMKIVTFALFPAIFEGMTNVGNQALIPALQDLNLDLINDMHTALA SSLGLDLPANATGMSKMGALGFELPIYLVFVNCVITWELFVSNGHEVNECEMFVST ILFALIVPTTLQMOALGITFGGIYAKELFGGVGNFNNPDLAAGRAFLFAVPAQIS GDTVMVAAGFSGATGALISOMSGOGALDHTVYGAITMDADAVGNLPGSGMEVSTLA ILLGAVIYFTRITAAWRIITAGWIGMATSTLENLIGSETNPMESMPWHHFFVLGPA LCMVRNATPVSASFNTGKRWMTGALIGVNAVILIRYVNPAYPCGMALILFANLFAI EYIVQANIKRRARTNC"

5540..6274

/gene="HI0167"

5540..6274

/gene="HI0167"

/note="similar to PID:663269 PID:663272 PID:893413 percent identity: 34.36; identified by sequence similarity; putative"

/codon_start=1

/transl_table=11

/product="NADH:ubiquinone oxidoreductase, Na translocating"

/protein_id="AAC21838.1"

/db_xref="GI:1573124"

/translation="NAKFNKDSVGGTILVLLSLVCSITVAGSAVWLKPADEOKLL DKONILNVAAGLQENTNVKETAKEIEPFVLDLAGEYQADDSQALPADADKAR IRSRSTTEVYLVKDOGOYVILPYGGLGMSVMVGYSDVDPGNTINGITVYVQIG ETPGIGGLEENPMWASLFGKKRIFDQHPALINIVGQAQODEHSIDGLGATLTGNG VQGTFTWMSKDSGFGYLERLHSGAN"

6274..6900

/gene="HI0168"

6274..6900

/gene="HI0168"

/note="similar to PID:663269 PID:663273 PID:893414 percent identity: 74.88; identified by sequence similarity; putative"

/codon_start=1

/transl_table=11

/product="NADH:ubiquinone oxidoreductase, Na translocating"

/protein_id="AAC21839.1"

/db_xref="GI:1573125"

/translation="MSGKTSYKOLLAPLAKNNPDLAQLIGCSALAVTRKLEAFVNA ALAVTLVYTGSLNLFVILIRNYIPNSIRIIVQDIIISLVIVVQDILKAAVYGLSKOIS VEVGLITGNCIYVGRAEAFAMKSPVESFYDGLGNGISGSMVLIIVAFPEFLIGSGKL FGMTIFETIONGSGWQANGFLAPSAFIIIGVYVINGLRTWKPQGEK"

6904..7500

/gene="HI0170"

CDS

6904..7500
 /gene="H10170"
 /note="similar to PID:893416 PID:893415 percent identity: 84.34; identified by sequence similarity; putative"
 /codon_start=-1
 /transl_table=11
 /product="NADH:ubiquinone oxidoreductase, Na
 translocating"
 /protein_id="AAC21840.1"
 /db_xref="GI:1573126"
 /translation="MEHVISLFVKAFVLENNALSPFLGMCFTFLANSKVSPAPFGICIA
 VPEFLGIAVPVNOILITANYIKENALISGVDSIFLFTFISGLVLEIVLDFEM
 PSIIYALGIFLPIAVNCAIEGVSEFWORNDNPESTIVYEGSLGMMALVALAGL
 TERKRYADIPAGIKLGITFEISVGLMALGFMSFSGIOL"

Query Match 39.4%; Score 530; DB 1; Length 13299;
 Best Local Similarity 62.7%; Pred. No. 1.7e-118;
 Matches 844; Conservative 0; Mismatches 497; Indels 6; Gaps 2;

1 atgattaaatcaaaaaaggtctaaacctgcccacatcgcgagcagacgaagcattc 60
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 2966 ATGATTAAATTAAGAAAGTTTGATCTCCCAATTCGGGGAACCAACAGCAATATC 3025
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 61 tatgaagcgccgctcatcaccgaagtcgctgttgcttgaggaagaatacgcgatg 120
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 3026 CATAGCGGCAACCGTGTGATCAAGTTGCCATCTAGGTGAGAGATATGTGGGATGCG 3085
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 121 cccctngatgaagtcgaagaagcgatgcgctcaaaaaagcgaagtcgtttgaagac 180
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 3086 CCTTCATATGAAGTTCGCGAAGCGATGTTGTGAAAAAGGTCAAGTACTTTTGAAGAC 3145
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 181 aaaaagatccggcgctgtgttaccgcgcngtttcaagcaaatccgcgcacatc 240
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 3146 AAAAAAATCCGTGTGATTTTACAGCCCTCGCAAGCGGTACATCAGTCATCAT 3205
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 241 cgcgcgaaagcgctactcagtcgctgtgattcgcglttgaagcagcaaatc 300
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 3206 CCGCGGAAAAAGGTATTACATCTGTGTTTATATGAGAAAGGTGATGAAAAAATC 3265
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 301 gagtcgaacgctacgcgcgcgaagcgltgcaactaaagcgcgcgcgaatnngngc 360
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 3266 ACTTTTGGAAAAATAGACAGAAACATTTGATACACTTTTTCGAAACAACTAAAAACA 3325
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 361 aactcgatccaatccggttctgtgactgcgctgcgtanccgctccgltcagaanaatcc 420
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 3326 AACTGATGAATCTGGCTATAGACGGCATTTACGTACTGTCATTTAGTAAAGTTCCA 3385
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 421 gccgcgacgcgcgcgcgttcgcactcttcgtcaatcgatgagacacacacgcgtc 480
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 3386 TCCATTGAAGTGAAGCATCTTCTATTTTGTAAATCGGATGATACCAATCCATTACCA 3445
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 481 gcaagccctgtgtgtgtgacaaagacgcngcngangattcagaacgcnfngtctgt 540
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 3446 GCCGATCCCTCGTGTCTATTTAAAGATATTCGCAACACTTCAACGGTTTAAACCGTA 3505
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 541 ttgagcgcgttgcgcgcgctaa---aatccatgtgtgtaagcagctgcgcgcgcgt 597
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 3506 TTAAGTCGCTATTTCCTCAAAACCATTTACCTATGTATGAAGCAGCGGATCTTAATT 3565
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 598 ccgctgtaaaatgcctgcacacatcgaaacatgaattcgcgccgcgcgcgcgcgcgt 657
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 3566 CCAACTGCTGATCTTGAATAATTACAGATTCAATGATTTTACTGTTGTCATCCACTGCT 3625
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 658 ttgagtcgacacatcatcattcattcagccgctgcgtgcgaacaaacgcttggagc 717
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 3626 CTGTGAGGAGCCATATTCATTATTTGATCTGTCGCTATTTCAAAAAACGTATGGCAT 3685
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 718 atcaattacaagatgaattcgcatcgcgcgttctgttgcacagcgcgtctgaacacc 777
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 3686 ATCAATTATCAAGATGTGATCGCTGTAGGTAATTTATCAACACGCAAGCATTTATATCA 3745
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
 778 gagcgcgcatgcttgggtgttctcaagtaacaaacacgccttgcgtacgct 837
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

Db 3746 GAACGATAATTTCTCTTSCAGGTCTCTCAAGTGAAGAACCTCGTTAGTACGTACAACA 3805

Qy 838 ttggttggaagatctcgaaattactcgcgcggaatttggttaacgcagaacccgctg 897
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

Db 3806 ATCGGTCTAATCTTTCATAATTACCCAAATGAATTAAGTCAAGTAAATCCGTGA 3865

Qy 898 attcgcgtcgatctgaacgcgcgcgcgtatcacaaagcgcgcgcgcgtatcttggagc 957
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

Db 3866 ATTCTGGCTCAGTCTCTTGTGTGTCATAATAGCAAAAGATTACATGATTTTAGTCG 3925

Qy 958 taccacaatcgcattccgltatcgaagaagccgcgaagaagaagctgttcgcctgggt 1017
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

Db 3926 TATGCTTTGCAAGTATTCAGTAATTCCTGAAGGTAATGAAGAAATTTTGGTGGAT 3985

Qy 1018 ggc 1077
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

Db 3986 ATGCCACAGCCAAATTAATTTCTGTGATCGTACCGTATTAGGCCACATTC---AGTAAA 4042

Qy 1078 aaactctcaagtcacgcacgc 1137
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

Db 4043 AAATTTGTTAATTTTACTACTTCAGAAAAATGTGGCGAAGCTGCAATGGTCCATGCGT 4102

Qy 1138 acttaagc 1197
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

Db 4103 AGCTATGAGCGCGTAAATGCAATTGACATTTTACATTAATTTATACGTGATTAATC 4162

Qy 1198 gtgc 1257
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

Db 4163 GTGGGCGATACCGATGATGCTCAAGAGTGAAGTGTCTGTGAATTAACGAAAGACATTG 4222

Qy 1258 gcttctgcaactcgcgtcgc 1317
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

Db 4223 GCGTATGTTGCTTTCGTTGTCACAGCAAAATATGATATGAGTTCAATCTTTCGTCAGATC 4282

Qy 1318 ctggaacnltgsgaagagcgtga 1344
 ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||

Db 4283 TTAGTAAATTTGAGAAAGAGTTAA 4309

RESULT 10

AF165980 6048 bp DNA linear BCT 13-NOV-2001

LOCUS

DEFINITION

Vibrio harveyi Na⁺-translocating NADH-quinone oxidoreductase

ACCESSION

AF165980

VERSION

AF165980.1 GI:6531420

KEYWORDS

Na⁺-translocating NADH: ubiquinone oxidoreductase from Vibrio harveyi

SOURCE

Vibrio harveyi

ORGANISM

Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

REFERENCE

1 (bases 1 to 6048)

Authors Zhou, W., Bertsova, Y.V., Feng, B., Tsatsos, P., Verkhovskaya, M.L., Gennits, R.B., Bogachev, A.V. and Barginera, B.

Sequencing and preliminary characterization of the Na⁺-translocating NADH: ubiquinone oxidoreductase from Vibrio harveyi

JOURNAL

Biochemistry 38 (49), 16246-16252 (1999)

MEDLINE

20056044

PUBMED

10587447

REFERENCE

2 (bases 1 to 6048)

Authors Zhou, W., Feng, B., Gennits, R.B. and Barginera, B.

Direct Submission

Submitted (06-JUL-1999) Biochemistry, University of Illinois, 600 South Mathews Avenue, Urbana, IL 61801, USA

FEATURES

source

1..6048

/organism="Vibrio harveyi"

/strain="Bb120"

/db_xref="taxon:669"

169..1509

/gene="nqrA"

169..6022

/gene="Na⁺-translocating NADH-quinone oxidoreductase"

gene
CDS
complex operon"
169..1509
/gene="ngrb"
/codon_start=1
/transl_table=11
/product="NgrA"
/protein_id="AAF15411.1"
/db_xref="GI:6531421"
/translation="MIIKKGLAPLAPASOVINDKTIKKVALLGEEYVGMKRP
VRGDEVKRAQVLPEDKKNPVKFTAPAGKVIENRGA RVLQSVVIEVAGEEOV
DKFEAOLAGDREVITKQVLESGLMTALRTPESKVPALIESATKALFVAMPNPLA
AOPELITHEOQAEVAFAGLIDILSALPEGKVVYKSGSTLPSSQSNVEHEPDPHPG
LAGTHMILPYRNAENVMKSTIKOVYIARGOLFEGELTYDPRVSLAGPVPNPLRL
TVGASLEDLIDSELMPEEVRISSVLSGTASCPHAILGRTHOVSVIREGRDKL
FGMATPKNKRSTISFSGHIFKGLFNMFTTNGSDRAVP1GNETERVP1LMEPEPL
LRLDAGDTDSAQTLGALDELDEDLALCTCPGKRYEGQLRECLITVKBG"
1513..2754
/gene="ngrb"
1513..2754
/gene="ngrb"
/codon_start=1
/transl_table=11
/product="NgrB"
/protein_id="AAF15412.1"
/db_xref="GI:6531422"
/translation="MKRTIEDIEHHPERGGRHEKWFALYEAAATLFYTPGIVTKRSS
LHRBSVDLRIMVWFVAFPAFWGMVINGOALIALNMHITAGDOLATYIAGMHW
LHEMGGTIGAGAGVSMMLGATVPLPYATVIVGGEWEVLECMVRKHEVNGEFV
TSLFALLVPTPLMQAALGITFGVVAKEIFGCGRMFLMADARLFFAYPQ
ISGVYMTADEFSGATLSOMAHRGSGALINNTGAPITMMAAFIGNIGSIEVST
LALMGAMIVYRIASMRITAGVMGMLAVSLFNVIGSDTDMEMPMHMLVLSG
PARGFMATDPVSASFSTNSGKMWYICILIGAMCMIRVNPALPEGMMLALIFRNLFA
PLFDHVIERKNTKRLARIYK"
2744..3529
/gene="ngrc"
2744..3529
/gene="ngrc"
/codon_start=1
/transl_table=11
/product="NgrC"
/protein_id="AAF15413.1"
/db_xref="GI:6531423"
/translation="MASNDSIKRTLGVVIGLSVCSIIIVSTAAGLRDQOKANAVLD
KOSKIVEAGIDAEKKVPELFAEYIEPRDPTGSEVFEKAEKGSAAANDOKRAK
DPAESIKLTADDEKAKILIRANNGIYLVKNGDLSKVIIPVHGNGISMVYAVAVE
TGNNTVSGITTYEQGETPGIGEVBNPVRAPYKGLFEDENKPAIKIYKGAPEGS
EHGVDELGSATLITNGVQGTDFWLDGMGFPLAKVDDGLN"
3529..4161
/gene="ngrc"
3529..4161
/gene="ngrc"
/codon_start=1
/transl_table=11
/product="NgrD"
/protein_id="AAF15414.1"
/db_xref="GI:6531424"
/translation="MSSAONIKKSIAAPVUDNPALOVIGCSALAVTTKLEAFV
TLAVITVALSNFSVSLIRNHIPNSVRIYOMATIASIYIVDYVLAKEYLIDISKQLS
VFVGLITINCIVMGRAEAFAMKSAVPVSLIDGIGNGLGFGVLLITVGEFRELFGSKL
FGMEVPLVNSNGMYQPNGLMLAPSAFLIGFLIWIIRIKIPROVEAKE"
4169..4765
/gene="ngre"
4169..4765
/gene="ngre"
/codon_start=1
/transl_table=11
/product="Ngre"
/protein_id="AAF15415.1"
/db_xref="GI:6531425"
/translation="MEHYISLIYKSTIFENMALSPFLCMCFPLASKRVKTSFGVGA
VVVLTIAVENVNLVNLVLEKNALVEGVDLSFTNITFTFVIALVOILEKVDLRF
PPLYNALGIFLPLITVNCALFGVSPVQVRODYNAESLIVFGSGVGMMLAIVALGI
REMKYSDVPGRLGITEITVGLMALGEMSPSGVOI"

gene
CDS
4799..6022
/gene="ngrf"
4799..6022
/gene="ngrf"
/codon_start=1
/transl_table=11
/product="NgrF"
/protein_id="AAF15416.1"
/db_xref="GI:6531426"
/translation="WDILGVVMTFLIVLALVYLIFAKSKLVPFGDITISVNDPSL
AIVTPGKRLSALAGVAVSVSAGGSGCQCRKRTKSGGDLIPDLDTTKEA
REGERLACQVMKEDMDIELPEIEGVKMECTVINSNDKATFIELKQIIPGSEVP
FRAGYIQIEAPAHVVKYADYIPEIEYEDMKFNLFRYSKVNERTIYASVANEPE
EHGIIMLVRIATPEPNNDVAPGIMSSPILSKRDKCTISGPEEPKADDAEM
FVGGAGAPAPRSHTEPDOLKRLHRSKRXSPFWGARSKRMFVDEPDGLAADNFMV
HCALSDPLPEDMDGTYGTIHNVLENTLRDHEAPEDCEYIMCGPPMAAAVYIMLKD
LGVEDENILDDFEG"
BASE COUNT 1571 a 1351 c 1455 g 1671 t
ORIGIN
Query Match 37.7%; Score 506.2; DB 1; Length 6048;
Best Local Similarity 61.1%; Pred. No. 1.1e-112;
Matches 821; Conservative 0; Mismatches 520; Indels 3; Gaps 1;
QY 1 atgattaaatacaaaaaggtctaaccctgcacccatcgccgagccgagcaagcaatcatt 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 169 atgattacataaagaagggctttggctcttctatcgcagagacctccatccagatgatt 228
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 tatgacggcccgctacataccgaagtcgcttgcttgccgaagaataatgcggtatgcgc 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 229 aatgatgataagaccatcaaaaaagtcgcttggcgaagatcggttgacatcgct 288
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 cccctngataaagatcaagaagcgatgcggtcaaaaaagccaaagtcggttgaaagac 180
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 289 cctaccatgcattcgcgcgttgccgatgaataaaaaagccagattctttttgaagac 348
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 aaaaagatccgagcgctggtgattaccgcgcggttccaagcaaaatccgcgcatccat 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 349 aaaaagaccctgcgcttaagttacatgcctccagaccgctgaagatcgaaattaac 408
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 ccgcgcgaagaagcgcgtaactcaagtcggtcggtatgcggttggaagcaagcaaatc 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 409 cgtggcgctaacgctgacttcaatcagttgcttaagatgaagcgaagacaggtt 468
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 gagttgaagc 360
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 469 acgtttgatragtttcgaagcagctcaactcagagcttagatccggaagtatcaagact 528
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 aatcgtacatccatcggttctgtagtgcgtgtaaccgcttcggttcgaagcaaatccct 420
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 529 caattggatgaattcgcccttggaccgctttacgtactgcgtccgctcagcaagcttcca 588
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 421 gccgctgtagtcgaagccgcttcgccaacttcgtcaatgcgtagcaaccaatccgctngc 480
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 589 gcaatgcagctctgctactaagagctattttgctaactcgagtgatgataaactcaattaca 648
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 481 gcagacccttgcttgtagtcaaaagacccgnganganttcagacganganltcgtgta 540
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 649 gcacagccttgatgatacattcgaacaacagacatttcgttgacagctctacacatt 708
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 541 ttgagccgcttcgagcgaaataatccatctgltgaaggaagtcgagcgcgcgcgcgcgcgc 600
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 709 ctttcagcccttgacagagcaagcattttagctatgtaag---tcggcactagcttgcca 765
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 601 tctgaanaatgctgcaacatcgaacaacatgaattcggcgccgcgcgcgcgcgcgcgcgcgc 660
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 766 agctcttctcagtcgaacgtagaagacacgtcttcagtgtcttcacaccgacgctgtt 825
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 661 agtggcagcagcatcttaattcattgagccgctgcgttgcaacaaacgcgtttggaacatc 720
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 826 gctggcaccacacatgacttttccctataccacgtaagtgcgaanaattgtgcgctgagcattc 885
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||


```
gene
CDS
    /protein_id="AAD29966.1"
    /db_xref="GI:4808846"
    /translation="MEHVISLVKSFIFENMALSPFLGQTFPLAKSKVTSFGSLGIA
    VIVLVITVSPVNNVLYNNVILKPDALVEGVDLSPLNITTFGVAIVLLEMLIDRF
    PLVNALGIFLPLITVNNCAIFEGVSFMYVRDVSFAISVYVFGSGVMALAIYALAGI
    REKMYSDVPELPRLGIGITFTAGLMALFMSFSVOL"
    /gene="ngfr"
    /gene_synonym="4657..5883"
    /note="beta subunit of NQR enzyme complex"
    /codon_start=1
    /transl_table=11
    /product="Ngfr"
    /protein_id="AAD29967.1"
    /db_xref="GI:4808847"
    /translation="MSTIRFYVMFLIILALVYLIFPKSKVLPEDITISINGDE
    KATVDPGSKLTLLALAGAVSVSSACGGSGSCQVVKIKSGGDLITPLDHSKE
    AREGRILACQVAVKADMDLEPEEIFGVKMECTVLSNKNKATFIELKLAIIPGESV
    PERAGGYIOIEAPAHVHKYADFDVEKYRDMDFNLFRKSEYDEIIRAYSMANP
    EEEGILMLNVRATPPNNPNVPGQMSYVNSLKGDKCTIGPFEFPKADDAEM
    VFIGGAGAPMRSHIPDOLKRLKSRKMSYWGASRKREMFVEDPDGLAENDNV
    MHCAISDPQEDNMTGYTFHNNVLENTLKHAEDECEYIYCGPPMNAAYINMLK
    NIGVEENTLLDDEG"
BASE COUNT      1441 a      1276 c      1550 g      1649 t
ORIGIN
Query Match      37.4%; Score 503; DB 1; Length 5916;
Best Local Similarity 60.9%; Pred. No. 6,6e-112;
Matches 819; Conservative 0; Mismatches 522; Indels 3; Gaps 1;
QY 1 atgattaatacaaaaagagctcaaacctgccatcgcgagcgagcgagcaagtcatt 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 34 ATGATTACATATAAAAAGGATTTGGACCTTCCTATCCAGGAACCTCCTCCAGAGTATT 93
QY 61 tatgacggcccgcatctaccgaagtcgctgtcttggcgagaagataatgcccgtatg 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 94 AGTGATGGTAAAGCCATCAAAAAGATCGCCTTGCTGGCGAAGAGTACGTTGGCATGGT 153
QY 121 cccctgataaagtcgaagaagcgatgcccgtcaaaaaagcccaagtcgttctgaagac 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 154 CCTACTATGTCATGTCGCGAGTGTGATGAATAAAAAGCTCAGATTCTTTTGAAAGAT 213
QY 181 aaaaagaatcccgctgtgttaccgcgccngttccaggaataatccgcccattccat 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 214 AAGAAATATCCGGGTGTGAATTTACTTCACCACTACGCGGTAAAGTGGTTGAAATCAAT 273
QY 241 cgcgcgcaaaaagcgctacttcagtcgctgtatgcttgaagcgcaacgcaaatc 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 274 CGTGGTCAAAAGCGGCTTCTTCAAGTGTGTGATTTGAAGTGGCGGGTGAATGCAAGTG 333
QY 301 gaattcgaaacgctacgcccgaagcgctggaacttaagcgcganganngnngc 360
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 334 ACCCTTGACAAAGTTTGAAGCAATTCATTAAGCCAGTTTGATTCGCGATGCGATAAAACT 393
QY 361 aatctatcaaatcccgctgtgtgactgctgttancgcttcgttccagaataatccct 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 394 CAGCTGTGAGTCTGTGCTTCTTGGACTGCTTTCGTACTCTCGCTTCAGGAAGGTGCCA 453
QY 421 gccctgcatgagcgagcgcttcgccatcttcgtcaatgcatggaacacaatcccgctngc 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 454 GCCATTGATTCACCTCTGTGAAGCATTTTGTCACTCTGTATGATACCAATCCGTTGGCA 513
QY 481 gcaagaccctgtgtgtatgataaagaagccngcngangattccaagcangntnctgtgta 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 514 GGTGACCCGACCCCTGTGATCAATGAACAATCGAAGCCCTTGCTGCTGATGATGTG 573
QY 541 ttgagccgttttgacgagcgtaaatccatgctgtgtgaagcagctgagcgcaagctgagc 600
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 574 CTTTCCGCACTGACAACTGGCAAAAGTGTACGTTTGTAA---AAAAGGACCAAGCTTGCCA 630
QY 601 tctgaaaatgctgcacaacacatgaattcgcgccgcatccgagcttgg 660
```

```
DB 631 CTTTCTCAACACCAGAAATGTTGAGAAACATGCTTTGACGGCCCAATCCTCGACGCTTA 690
QY 661 agtgcagcgacatcattcattctatgacgctgctgctgcaaaaacacccgtttgacatcc 720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 691 GCTGCACTCAATGATGATTTCTCTTATCCATGCAAGTAAAGTGAACACACGCTGGGTGAGCATC 750
QY 721 aattatcaagaatgtaattgcacatcgagcgtttgttgcaacagagcgctctgaacaccgag 780
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 751 AACTATACAGAGAGTGTATGTCAGTAGTCAACTTCTCTGACGGGCGAGCTATACACCCAG 810
QY 781 cgcgtgattgcttgggtgttctcgaagtcacaacacacacgctcttggctacgctttg 840
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 811 CTTGTGTTCTCTTCCCGCCCGCGGTGTGAACAAACCTGTTGGTTCACACTGTTCATG 870
QY 841 gttggaagaatgctgcaaatctactgagcggaattglttgacgcaacacacccgctgatt 900
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 871 GGCGCATCTCTGAGACACTGCTGTGACACGGAATATATGCTTGCCGAGTACGATCATC 930
QY 901 tccgcttcgctatctgaacgcgcgatatacaacagcgcgacgcatlalttggagcgtac 960
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 931 TCTGGCTCTGTGCTGTCTGTGTAAGCAAGCACTGGAACACAGCTTATCTTGGTCGTTAC 990
QY 961 caacaatcagattccgttctcgaagaagcgcgacgcaagaagcgtgttcgctgtgtgctg 1020
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 991 CACTTACAGAGTCTGCTGTTCTTCTGTAAGGACGTGATTAAGAACTGTTGGTGGCGATG 1050
QY 1021 ccgagccgagcaaaatcattccatccatcgcttagacacctgacattcccgtaaaaacaa 1080
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1051 CCGGGTAAACAAAGTTCTCTGTGACTCGCTCTTTCTTGTGTGTAACATTCAAGGCGCAA 1110
QY 1081 ccttcaaatctcaacgaacgctgcaacggttggcagccgcacatgctgacattgctact 1140
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1111 GGTACAAACATGACCAACCAACTAATGATGATGATGATGATGATGATGATGATGATGATG 1170
QY 1141 taagagcgctaatccgcttagacatccctgacacctgcttcttggcgagttatcgctc 1200
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1171 TATGAGAAAGTATGACCTAGATATGAGAACCACTGTTGCTGCTGTGATCTGTGTGCT 1230
QY 1201 ggcgaatccgagcagcgcgcaagaatggtgtgtgttgtaattggaagcaagaacacctcgt 1260
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1231 GGTGACAGTGAACGCGCGGTGCTTATGATGCTTTAGAGCTGGATGGAAGAAAGATCTAGCA 1290
QY 1261 ttgtgagcttcgctcgtccgagcaataacgaatlangcccgctgttgcgtaagtgctg 1320
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1291 CTGTGATCTGCTTTGCCAGGCAATATAGAGTACGCTCAGCTCCGCTGAGTGTGCTTA 1350
QY 1321 gaaacncttgagaagaagcgctga 1344
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1351 GACAAAGATCGAGAAAGAGGCTAA 1374
RESULT 12
AE004300/c 13382 bp DNA linear BCF 31-JUL-2000
LOCUS
DEFINITION
Vibrio cholerae chromosome I, section 208 of 251 of the complete
ACCESSION
AE004300 AE003852
VERSION
AE004300.1 GI:9656850
KEYWORDS
SOURCE
ORGANISM
Vibrio cholerae.
Vibrio cholerae
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE
1 (bases 1 to 13382)
Heldelberg,J.F., Eissen,J.A., Nelson,W.C., Clayton,R.A., Gwin,M.L.,
Dodson,R.J., Haft,D.H., Hickey,E.K., Peterson,J.D., Umayam,L.,
Gill,S.R., Nelson,K.E., Read,T.D., Tetteijn,H., Richardson,D.,
Ermolaeva,M.D., Yamathevan,J., Bae,S., Qin,H., Dracoli,I.,
Sellers,P., McDonald,L., Uterback,T., Fleischmann,R.D.,
Nierman,W.C. and White,O.
DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae
JOURNAL
Nature 406 (6795), 477-483 (2000)
```

MEDLINE
REFERENCE
AUTHORS
20406833
2 (bases 1 to 13382)
Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwin, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.A.,
Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
Ermlaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragol, I.,
Sellers, P., McDonald, L., Utterback, T., Fleischmann, R.D.,
Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,
Mekalanos, J.J., Venter, J.C. and Fraser, C.M.
TITLE
Direct Submission
Submitted (14-JUN-2000) The Institute for Genomic Research, 9712
JOURNAL
Medical Center Dr. Rockville, MD 20850, USA
FEATURES
Source
1. 13382
/organism="Vibrio cholerae"
/strain="N16961"
/serotype="O1"
/db_xref="taxon:666"
/chromosome="1"
/note="Biotype: El Tor"
/complement(154. .1236)
/gene="VC2287"
complement(154. .1236)
/gene="VC2287"
/note="similar to GB:000096 SP:Q47155 PID:1552799
PID:1786425; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="DNA-damage-inducible protein p"
/protein_id="AAF95431.1"
/db_xref="GI:9656851"
/translation="MODRIKIIVHDMDCFFAAVEMKNDPAVRIEALAVGHEKORGV
ISTCYQARKFVSAMPTAOLALCPOLHVVVGKMSVYKSSQIQIFQRYSLIE
PLSLDEAYLDVSESTAYOGSAPLIAQIRIDIMQELMTASGAVAFIKFLKVASDLN
KPDGLVYVPDQVQEMVDSLEPIKIPGVKAGLEKHOAGIYGVADYRRADYRKILHO
RFGALSLMKRSHGIDEREYMERKRSYVETESQNISITFCRCMOYITEKILPELD
ARLSRAHPQRGITKOSIKYKFAFDFOQTIEHVHPEALDIEVLTLEOVLTQQGHEIR
LIGLSVLMKPELQMKQLSMPPSDGMQ"
complement(1345. .1578)
/gene="VC2288"
complement(1345. .1578)
/note="VC2288"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF95432.1"
/db_xref="GI:9656852"
/translation="MNTFLITEGFLAVIAAAVGIIFQRKYVKGSCGGGLGAVGIEKY
CNCPPCDARRKREKAAKAEKRLAMEKDRIA"
complement(1595. .2698)
/gene="VC2289"
complement(1595. .2698)
/note="VC2289"
/note="similar to GB:I42023 SP:P44550 PID:1003244
PID:1222087 PID:1204428; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="thiamin biosynthesis lipoprotein ApbE"
/protein_id="AAP95433.1"
/db_xref="GI:9656853"
/translation="MEFTMWILKSVNMFVRSVFPISIFNYSKQVNRMLVALAST
LLAGCEKPAEQVHLSGPTMGTYNINIKYIQOQPIADSKTLOTEDRLIEVNDQMSY
KLSSELSPNGHTSESPFAVSTQTLTFVKEAIRLNEGALDVTGVLIVLMWEGPE
ARPDVYPMDEELNARRAIRTGIEHLITEGNTLSKDIPELVYDLSIAGCMGVADYIL
OSQGIENMYEIGETRLKGLNBDGYPWRAIKRPSYDORSVDPIEFGDAIATSGD
IRNTEQDGVRSIIIDPTTGRPIRNNRVSVTLDSKMTADGLATGLMWGEERGMA
VAEANQIVLMVKTDDGKFKYASSSEKPELSK"
complement(2745. .3971)
/gene="VC2290"
complement(2745. .3971)
/gene="VC2290"

/note="similar to PID:893416; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH:ubiquinone oxidoreductase, Na
translocating, beta subunit"
/protein_id="AAF95434.1"
/db_xref="GI:9656854"
/translation="MSTRIIFGVWFTLIIALVYITFAKSLVPGDITISINSGPE
KATVOPCKLITLALAGAVFVSSACGGGSGCGORVIRKSGGDIPLPELDHISGE
ARBERLACQAVAKADMDELPELIFGVKKECTVISDNKATFTELKLAIPDGESV
PFRAGYTIQLEAPAHVHKYADVDPEKRGMDKPLFRYESKVDPEIIRAYSMAYP
EEFGIIMLVNRIATPPNNPNVPPGOMSSYIWSLKAGDKCTISGPFEPFAQTDDEM
VFIFGCGAMAPMRSHIFDOLKRLKRSKMSYWCARSKREMFYEDPDGLAENDNFV
WICALSDPOEDNMWYTGTFIHVLYENYIMDHEAPECEYIMCGPPMNAVIMNLK
NLGVEENIILLDDRCG"
complement(4005. .4601)
/gene="VC2291"
complement(4005. .4601)
/gene="VC2291"
/note="similar to PID:893416 PID:893415; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH:ubiquinone oxidoreductase, Na
translocating, hydrophobic membrane protein NqrE"
/protein_id="AAF95435.1"
/db_xref="GI:9656855"
/translation="MERYTSLVKSIFENNALSFLLGMCFLAVSKKYSFFGLGIA
VIVVLTISVPNNLVNVLKPDALVEGVDSLFNFTFIVIALVQILEMILREF
PLLYNALGFLPLITVNCALFEGVGFVWQRYSPAESVYVGGSGVWMLAIVALLAI
REMKYSDVPGLGLGITFTTALMALGFMSFSGVQ"
complement(4609. .5241)
/gene="VC2292"
complement(4609. .5241)
/note="VC2292"
/note="similar to PID:663269 PID:663273 PID:893414;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH:ubiquinone oxidoreductase, Na
translocating, hydrophobic membrane protein NqrD"
/protein_id="AAF95436.1"
/db_xref="GI:9656856"
/translation="MSSAKELKRSVLAFLVDNNPIALGVLSALAVTTKLEFAFM
TLAVFVYATLSNFEVSLIRNIPNSVRIYOMATIASDVIYVDILKAVLYXISOLS
VFGILITNCTVWGRAPAPAKSPISQIDIGNGIGCYVLMVGPPELLSGKL
FGLLEVPLISNGCYQPNGLMLAPSAFLIGFMIAIRTRPEQVEAKE"
complement(5241. .6014)
/gene="VC2293"
complement(5241. .6014)
/note="VC2293"
/note="similar to PID:663269 PID:663272 PID:893413;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH:ubiquinone oxidoreductase, Na
translocating, gamma subunit"
/protein_id="AAP95437.1"
/db_xref="GI:9656857"
/translation="MASNDSIKTKLEFVIALSLVCSITVSAAYGLDOKOKENALD
KSKRILOVAGIEAKSKQIVLEFNKSLIEPLVDFTGPDVGDAAVYDOKRAAKASE
SILKLAEDQAKIQRANVGVVYLVKDGDKSVYLIPVHGVLMSMAYAFVETDGN
TVGGLIYVEGEGVGLGCEVNPAPRAQWVCKLDEENHKPAIKIVKGAAPGSGHGV
DGLSGATITISNGVONTDFMLGDMGFPFLIKVYDGLN"
complement(6001. .7284)
/gene="VC2294"
complement(6001. .7284)
/note="VC2294"
/note="similar to PID:663269 PID:663271; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11

[illegible][illegible]


```

Db      1168  CAGTTGGTTGACTCTGCGCCATATGACCGCTTACGTACTCGTCGGTTAGCAAGTTCT 1227
QY      421  gccgtcgatgacgagcggttcgcacatcttcgtcaatgcatgagacacaatccgctngc 480
Db      1228  GCATATGAGCTTCTACTAAGGCTATTTTCTTAACATGCAATGACTAATCAATTACCA 1287
QY      481  gcagaccctgtgtgtgatcaagaagccgncngangattcagacganganctcgtgta 540
Db      1288  GCNAACCTCAGTTGATCATTTAATGACCAACAAACACATTTCATTCGCTGTACACATT 1347
QY      541  ttgagccgttttgacccgagcgtlaaatccatgtytgaaggcagctggcgacagctgcg 600
Db      1348  CTTTCAGCCCTGACAGAGGACGAGGTTTACGTATGTAA--ATCTGGACATAGTTTGCCA 1404
QY      601  tctgaaaatgctgccacaatcgaacatgtaattcgcgccgacatccgcggtttg 660
Db      1405  CGCTCTTCAACAGTCTAAGCAGTAAGAAACACGCTCTTCATGAGCCCTCACACCGACGCTT 1464
QY      661  agtgcagcagacatcatcttcattgagcgtgctgctgcaacaacacgctttgacacatc 720
Db      1465  GCTGGACACCCATGTCATTTCTTATACCCAGTGAATGCAGAAAATGTGGCATGGAGCATC 1524
QY      721  aattacaagatgtaattgcatcgagcgtttgtttgcaacagcgctcgtgaacacgag 780
Db      1525  AACTACCAAGACGCTTATCCGCTCGTAAGCTGTTCCTAAGTGGTGAACCTATACAGCAT 1584
QY      781  cgcgtgatgtcttctgtgtgttctcaagtcaacaacacgctcttggctacgcttttg 840
Db      1585  CGGTGTGTGCTAGTGTAGTGTCCAGTTGTGATTAACCTCTCTAGTTCACCTTAATTC 1644
QY      841  ggtgcgaagatcgcgaatctactgctggcggaattggtttgacgagacacgcgctgatt 900
Db      1645  GGTGCAAGCTTGATGATGACCTTAACAGATACGAGATTGATGCTGGCGAAGTTCGTGTATT 1704
QY      901  tccggttcggtatltgaacgagcgatatacaacagcgcgacgattatltgagacgtac 960
Db      1705  TCTGTTCTGTACTGACAGGTACACATGCAACTGTCCACACGATACCTTGGCGCTTAC 1764
QY      961  caaatcagaattccgcttaccgaagaagcgcgacgaagcgcttgcgtgtgtgtcg 1020
Db      1765  CATCAACAGGTTTCTGTATTACGTGAAGCGCGGTGAGAAAGACTATTGCTGGGCAATG 1824
QY      1021  ccgacgacgagacaatactcatcagcgttaacgacctcgccatcttcctgaatacaaa 1080
Db      1825  CCGTGTAAGATAAAGTTCTCGGTAACTCGCTTCTTCCGTGTCACGTATTCAAAGCTCAG 1884
QY      1081  ctctcaagltcacgacagcgcgtcaacggtgagcgcgacgctgtgacgttgcattact 1140
Db      1885  TTGTTCAATATGACAAACGACGACCAATGATGATGATGATGATGATGATGATGATGATG 1944
QY      1141  taagaagcggtatagcgcgtcagacatcctgctacccctgcttctggcgatltatgctc 1200
Db      1945  TACGAACGCTTAATGCTCTGATATGAGATGAGCTCTGCTGCTGCTGCTGCTGATGATG 2004
QY      1201  ggcgataccgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacg 1260
Db      2005  GGTGATACGACACACTGCTCAAGCACTGGTCCGCTTGAACGTGATGAAGAAATATACGA 2064
QY      1261  ttgtgagactcgtctgcccgggcaataacgaatlangcccgctgtgtgcaagtgtcgt 1320
Db      2065  TTGTGTAACCTTTGTATGTCCAGGTAAGTACGATGAGTACTTCTGTAATGCTGA 2124
QY      1321  gaaacnttgaagaagaagctga 1344
Db      2125  GATTAATATGAGAGGAGGTAA 2148

```

RESULT 14

AB008030 6908 bp DNA circular BCT 15-JUN-1999
 LOCUS AB008030
 DEFINITION Vibrio alginolyticus genes for Na-translocating NADH-quinone reductase complex, nqr operon, complete genome.

ACCESSION

AB008030 D43958 D49364
 VERSION AB008030.1 GI:2558472
 KEYWORDS Nqr1 subunit of Na-translocating NADH-quinone reductase complex

SOURCE

Vibrio alginolyticus DNA.

ORGANISM

Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

REFERENCE

1 (sites)
 Hayashi, M., Hirai, K. and Unemoto, T.
 Cloning of the Na(+)-translocating NADH-quinone reductase gene from the marine bacterium Vibrio alginolyticus and the expression of the beta-subunit in Escherichia coli

AUTHORS

FEBS Lett. 356 (2-3), 330-332 (1994)

TITLE

2 (sites)
 Beatlie, P., Tan, K., Bourne, R.M., Leach, D., Rich, P.R. and Ward, F.B.
 Cloning and sequencing of four structural genes for the Na(+)-translocating NADH-ubiquinone oxidoreductase of Vibrio alginolyticus

AUTHORS

FEBS Lett. 356 (2-3), 333-338 (1994)

REFERENCE

3 (sites)
 Hayashi, M., Hirai, K. and Unemoto, T.
 Sequencing and the alignment of structural genes in the nqr operon encoding the Na(+)-translocating NADH-quinone reductase from Vibrio alginolyticus

AUTHORS

FEBS Lett. 356 (2-3), 333-338 (1994)

TITLE

4 (sites)
 Hayashi, M., Nakayama, Y. and Unemoto, T.
 Existence of Na+-translocating NADH-quinone reductase in Haemophilus influenzae

AUTHORS

FEBS Lett. 381 (3), 174-176 (1996)

REFERENCE

5 (sites)
 Nakayama, Y., Hayashi, M. and Unemoto, T.
 Identification of six subunits constituting Na+-translocating NADH-quinone reductase from the marine vibrio alginolyticus

AUTHORS

FEBS Lett. 422 (2), 240-242 (1998)

TITLE

6 (bases 1 to 6908)
 Hayashi, M., Unemoto, T. and Sugiyama, A.
 Direct Submission
 Submitted (09-Oct-1997) Mak1 Hayashi, Faculty of Pharmaceutical Sciences, Chiba University, Laboratory of Membrane Biochemistry; 1-33 Yayoi-cho, Inage-ku, Chiba, Chiba 263, Japan
 (E-mail:maki@phs.chiba-u.ac.jp, Tel:043-290-2932, Fax:043-290-3021)
 On Oct 22, 1997 this sequence version replaced gi:677955 gi:893412.
 D43958: submitted (17-Dec-1994)
 D49364: submitted (17-Feb-1995).
 Location/Qualifiers

FEATURES

1..6908
 /organism="Vibrio alginolyticus"
 /db_xref="taxon:663"
 /note="There are conflicts between this sequence and GenBank Accession Number:Z37111."
 386..1726
 /gene="nqr1"
 386..1726
 /gene="nqr1"
 /note="contains FMN as a cofactor; nqrA, Nqr1 subunit of Na-NOR complex alpha-subunit"
 /codon_start=1
 /transl_table=11
 /evidence=experimental
 /product="Nqr1 subunit of Na-translocating NADH-quinone reductase complex alpha-subunit"
 /protein_id="BAA22910.1"

COMMENT

COMMENT

JOURNAL

JOURNAL

MEDLINE

MEDLINE

REFERENCE

REFERENCE

AUTHORS

AUTHORS

TITLE

TITLE

JOURNAL

JOURNAL

MEDLINE

MEDLINE

REFERENCE

REFERENCE

AUTHORS

AUTHORS

TITLE

TITLE

JOURNAL

JOURNAL

MEDLINE

MEDLINE

REFERENCE

REFERENCE

AUTHORS

AUTHORS

TITLE

TITLE

JOURNAL

JOURNAL

MEDLINE

MEDLINE

REFERENCE

REFERENCE

AUTHORS

AUTHORS

TITLE

TITLE

JOURNAL

JOURNAL

[illegible]

Db	15550	GGCGACACAGGAATG6GCCAAGAGCTGGCGCCCTCGGAACCTGCAGACGAAGAGACTTGGCG	15491
Qy	1261	ttgtgcagcttcctgcctgcgcggcgaataacgaataangccgcgtgtgtcgtlaagtgctg	1320
Db	15490	CTGTGCACCTACGTCTTGTGGCGCGCAAGTACGAAATFACGGCCGATTCCTCGGGACCAATCTG	15431
Qy	1321	gaaaccttggaaaggaagcttga	1344
Db	15430	GCACGCATCGACGACGAGGCGCTGA	15407
RESULT	16		
LOCUS	A96090	474 bp	DNA
DEFINITION	Sequence 123 from Patent WO924578.		linear
ACCESSION	A96090		
VERSION	A96090.1	GI:6779937	
KEYWORDS	unidentified.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 474)		
AUTHORS	Pizza, M., Scariato, V., Rappuoli, R., Grandi, G. and Masiugnani, V.		
TITLE	Neisserial antigens		
JOURNAL	Patent: WO 9924578-A 123 20-MAY-1999; PIZZA MARIAGRAZIA (IT); SCARIATO VINCENZO (IT); RAPPUOLI RINO (IT); CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)		
FEATURES	location/Qualifiers		
source	1..474		
BASE COUNT	122 a 130 c 130 g 91 t		1 others
ORIGIN			
Query Match	32.1%; Score 431.6; DB 6; Length 474;		
Best Local Similarity	93.5%; Pred. No. 1.6e-94;		
Matches	443; Conservative 1; Mismatches 30; Indels 0; Gaps 0;		
Qy	1	atgattaaatcaaaaaagtgtaaacctgcgccacgcgcggagacgcgagcaagtcatt	60
Db	1	ATGATTAAATCAAAAAAGTGTAAACCTGCCCATCGCGGACGACCGAGAACCGTT	60
Qy	61	tatgacgagcccgctacatcaccgaagtcgcgtgtccttggcgaagaatatgcggtatgcgc	120
Db	61	TACGAGCGCGCGCCGTCATTACCGAAGTCGGGTGCTTGGGAAGATATGCCGTATGCGC	120
Qy	121	cctcgtatgaagtcgaaggaagcgatgcgcgttcaaaaaagccaaagtctgttlyaaagc	180
Db	121	CCCTCGATGAAGTCAAGGAMGCGCATGCGCTCAAAAAAGCCAAAGTCTGTTTAAAGAC	180
Qy	181	aaaaagatccgggggtgtgttcttcgcggcgnfllttaaggaataatccgcgcacatc	240
Db	181	AAAAAGAAATCCGGGGTGTGTGTTCTGTGGCGGCTTCAAGGAAATATGCGCGATTCAC	240
Qy	241	cgcgcgaaaagcgcgtacttcagtcgcgtgcgtatgcgcgttgaagagcgaacgacaaalc	300
Db	241	CGTGGCGAAAACGCCGTACTTTCAGTCAGTCGATGTCGGTTGAARGCACAGCAAAATC	300
Qy	301	gaattcgaacgctacgcgcgcgaagcgttbgcaacttaagcgcgcganganatnnngmc	360
Db	301	GAGTTTGAACGCTACGCACTGAAGCGGTGCGAAACTTAAGCGGCGAAGAAATGTCGCCGC	360
Qy	361	aatctgatacaatccggttctgtggaatcgcgctgtanccgctcgttcaagcaaatccct	420
Db	361	AACCTGATCAATCCGTTTGTGGATGTGGCGCGCACACCCTGCTTCAAGCAAAATTCCT	420
Qy	421	gacgtcgaatgcgcgcgcttgcacatctctcftcaatgcgaatgcgaacccaatccg	474
Db	421	GCGCTGATGCCGAGCGCTTCCGCACTTTCGTCAATTCGATGGAACCAATCCG	474
RESULT	17		

VERSION D78193.1 GI:1064780
KEYWORDS phospho-beta-glucosidase; 6-phosphogluconate dehydrogenase; alkyl hydroperoxide reductase; amino acid permease; arginase; ornithine aminotransferase; positive regulatory protein; gntZ; ahp; bglA; yyc; yyc; formaldehyde dehydrogenase.
SOURCE Bacillus subtilis (strain:168) DNA.
ORGANISM Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Bacillus/Staphylococcus group; Bacillus.
REFERENCE 1 (bases 1 to 36200)
AUTHORS Ogasawara,N
TITLE Direct Submission
JOURNAL Submitted (06-NOV-1995) Naotake Ogasawara, Nara Institute of Science and Technology, Graduate School of Biological Sciences; 8916-5 Takayama-cho, Ikoma, Nara 630-01, Japan
MEDLINE (E-mail:nogasawads.aist-nara.ac.jp, Tel:07437-2-5430, Fax:07437-2-5439)
REFERENCE 2 (sites)
AUTHORS Zhang,J. and Aronson,A.
TITLE A Bacillus subtilis bglA gene encoding phospho-beta-glucosidase is inducible and closely linked to a NADH dehydrogenase-encoding gene
JOURNAL Gene 140 (1), 85-90 (1994)
MEDLINE 94171085
REFERENCE 3 (sites)
AUTHORS Hartford,O.M. and Dowds,B.C.
TITLE Isolation and characterization of a hydrogen peroxide resistant mutant of Bacillus subtilis
JOURNAL Microbiology 140 (Pt 2), 297-304 (1994)
MEDLINE 94236234
REFERENCE 4 (sites)
AUTHORS Ogasawara,N., Nakai,S. and Yoshikawa,H.
TITLE Systematic sequencing of the 180 kilobase region of the Bacillus subtilis chromosome containing the replication origin
JOURNAL DNA Res. 1 (1), 1-14 (1994)
MEDLINE 96051385
REFERENCE 5 (sites)
AUTHORS Calogero,S., Gardan,R., Glaeser,P., Schweizer,J., Rapoport,G. and Debarbouille,M.
TITLE RocR, a novel regulatory protein controlling arginine utilization in Bacillus subtilis, belongs to the NtrC/NtrA family of transcriptional activators
JOURNAL J. Bacteriol. 176 (5), 1234-1241 (1994)
MEDLINE 94156824
REFERENCE 6 (sites)
AUTHORS Yoshida,K., Seki,S., Fujimura,M., Miwa,Y. and Fujita,Y.
TITLE Cloning and sequencing of a 36-kb region of the Bacillus subtilis genome between the gnt and lol operons
JOURNAL DNA Res. 2 (2), 61-63 (1995)
MEDLINE 96093926
REFERENCE 7 (sites)
AUTHORS Gardan,R., Rapoport,G. and Debarbouille,M.
TITLE Expression of the rocDEF operon involved in arginine catabolism in Bacillus subtilis
JOURNAL J Mol Biol. 249 (5), 843-856 (1995)
MEDLINE 95311309
REFERENCE 8 (bases 1 to 36200)
AUTHORS Kashahara,Y., Nakai,S., Yoshikawa,H. and Ogasawara,N.
TITLE 36kb sequence between gntZ and trnY of B. subtilis genome Unpublished (1997)
FEATURES
Location/Qualifiers
1..36200
/organism="Bacillus subtilis"
/strain="168"
/db_xref="taxon:1423"
1..32
/gene="gntZ"
/gene="gntZ"
/gene="gntZ"
/gene="gntZ"
/citation=[6]
/codon_start=3
/transl_table=11
/product="6-phosphogluconate dehydrogenase"
/protein_id="BA11267.1"

gene
CDS
/db_xref="GI:1064781"
/translation="GVFHTNWESE"
496..1059
/gene="ahpC"
496..1059
/gene="ahpC"
/note="induced by hydrogen peroxide"
/citation=[3]
/codon_start=1
/transl_table=11
/product="alkyl hydroperoxide reductase"
/protein_id="BA11268.1"
/db_xref="GI:1064782"
/translation="MSLIGKEVLPFEAKAFKNGEFDVTNEDLKGMVFCFPADFS
FVCPLEBLOQVYALKEIGVEYVSSTDTPEVHKMGDSSEKISKITYAMIGDPSO
TISRNPVDVDEGTGLADRGFTIIDPGVIGQVEINNGSIGRASNLVNKVKAQYVRQ
NPGVCPAKMEGEGETLTPSLVNGKI"
1073..2602
/gene="ahpF"
1073..2602
/gene="ahpF"
/note="induced by hydrogen peroxide"
/citation=[3]
/codon_start=1
/transl_table=11
/product="alkyl hydroperoxide reductase"
/protein_id="BA11269.1"
/db_xref="GI:1064783"
/translation="WVLDNRIKAQLNQYQLIENDIVLKVSGEDTSDMLALVDEL
ASMSSKISVEKELNRPBESVNRVGEDEVGAVFAGIPDLHEFTSLVALLOVSGRPK
VQKVIDYQKIKSGEYHEPESYISLCHNCPDVVOALNMSVNPNTHTMIDGAAYKA
EVESKNIAVPYVLYNGESGSGRMTEELILAKMGSTGASPADKPEPDVAVGGP
AGASAATYARKGIRGVYAEERGGVLDITMSIENTSVKATGEPYLAISIEHVEAY
DIDVNIQIARAKIRKKDLFEELFENGCAVTKSTYILTSGARBRNVVPEQKNGCV
AYCPHCDGPLEGKDVAVAYGCGNSGTEAIDLAGVNHVTVLEFAPELADEVLQKRL
YSLPVTYVKNQAQKEITGDSQVNGITVYDRETEGKHEVLEQGFVQIGLVPTNTELE
GTVERNRGEIIVDHGATSVPEGLFAAGCTDSAVNQIILISMSGATPAALGAFVYLIR
N"
complement(2712..4151)
/gene="bglA"
complement(2712..4151)
/gene="bglA"
/citation=[2]
/codon_start=1
/transl_table=11
/product="phospho-beta-glucosidase"
/protein_id="BA11270.1"
/db_xref="GI:1064784"
/translation="MGNMPKDFLWGGALAAHQFEGMNGGKSPVVDVATGAGGVP
RKITDPIEENEFYPMHEAIDFVHRYKEDIALFAEMGLKLPISIGSRIFPGDEAF
NEAGLOFDDVDDELKGIPEVITLSHEMPHLHAREVGRNKKVDFPVFAFAC
FTRKAKDKYMMTENEINNMVNPNPLMTNSGYVGENAKKEMVQDTAHHELVAC
ALAYAKGDINPEROICAMVSHVPIPPSPNEDVMALEBKORIRFPFDVQKGYTP
SYALKEFEREGNITFEDGDDILRNGVVDYIGFSTYMSKTVKSDVKNNTGDIWGG
LPNGVENPYITSSDGMWAIDPGLVITLNRFDYRQIPLFIWENGAVADTLEEDSKY
HDPENIQYLIKSHIEALKRAVYDGVLDLIGYTPWGIIDIVSFTTGEMKKRYGMIVYDRD
NEGMSKRYKDSPEWYKNVYQTNGEEL"
4733..5449
/gene="yydK"
4733..5449
/gene="yydK"
/codon_start=1
/transl_table=11
/protein_id="BA11271.1"
/db_xref="GI:1064785"
/translation="WRMLKYOQIATEIETIYIEHLOQGDKPLVLETTMAQPEVSKST
ITKSLIELEQKGALEQVRSGLFVRKHKRGYISLSNQGFKLEDFNVNTSYIELD
VRKPTPEAENMINIGMEDIDYVVRVIRNGYISLSNQGFKLEDFNVNTSYIELD
FHYIREGIGLKIGESDPLFHWQOLNEEAEYVIGLBAQPKLYTESIFHLTNGOPFDYS
KISYNEQSFVVOANSFL"
complement(5766..6488)
/gene="yydJ"
complement(5766..6488)

```

/gene="yydJ"
/codon_start=1
/transl_table=11
/protein_id="BAA11272.1"
/db_xref="GI:1064786"
/translation="MKLEFKSISNKNVILLGAMFVFLFLGYFLVLDGKSVNTPTE
KCFPSSTVATQFGLMFSVIAFPIREFSNKNILEFKLIGENIYFFYKRIAVLF
MEFAFTILGLTISLMYHDSHFALLFELREYSAVILQYILIGTISVLCPNLISIGS
IYVMTSVILVAISNKPFIAPFEAGNIMYPIERLYLOSDNMLGSDNYLITLILY
SIITNAIVRSKTRMKKGL
complement(6509..7138)
/gene="yydI"
/complement(6509..7138)
/gene="yydH"
/codon_start=1
/transl_table=11
/protein_id="BAA11273.1"
/db_xref="GI:1064787"
/translation="NNIANYTLKVKGKTLIDQDPLHFSGKINHVGNKNGKSQLAK
DELNNSKRIGRIDRONSLISSSNIPNDVSKDELHPLSKREDAMIDKIVALLN
DNIDCKVLIKNLSGOKOKILISFLIEDTNIIVLDETISLDKKTIEIHLGFLNKYI
QENPEIITINITHDSLDKAIREDYITFNHQLQOYHSVDKLEIVTNE"
complement(7288..8046)
/gene="yydH"
/complement(7288..8046)
/gene="yydH"
/codon_start=1
/transl_table=11
/protein_id="BAA11274.1"
/db_xref="GI:1064788"
/translation="MKILKEDCKEYEVLVONNVFIDKKSGEYKNSLNSDKQLLR
FKMYKEVSPKPEYFLFLKFLKELMELIKPEYKQITLISNGFMGSKKVGQVYH
IYLHEGHYLSLKEFGKNEDKVGKFLNFYPAFYVOLNEMYSRREKIIVHLPGLF
IYLLNTLEILINQITFESSPALTMARPLESTLIMNLPIILNSGYKILLFLSLDER
SRKTNHMLVLTITQITIGLAVNSVHMITYN"
complement(8027..8986)
/gene="yydG"
/complement(8027..8986)
/gene="yydG"
/complement(8027..8986)
/gene="yydG"
/codon_start=1
/transl_table=11
/protein_id="BAA11275.1"
/db_xref="GI:1064789"
/translation="MYNKTVSINDSRCSNACDHCPSPTSTTRMEKEYIRLVE
FKNKTIQVISPFGVEFLDYKFLKELMELIKPEYKQITLISNGFMGSKKVGQVYH
DNNSINVLATISYDEVHAPFVSSSTKINLESRKYPDIDISNMAVTRKMSNHL
BELGDSILGKTIKFPWISVGAAKTRIKQENIHKFTYLEDSDJACHGQYDIVYHHDGE
ITPCSPALFETKITLREYVNSQSFERTVLSNLFILKKEGKFWFLNLIKNNKI
EEFDIPEFSSICGVSLENSAKINIFYPYMKYNNENKVV"
complement(9044..9193)

```

gene

Query Match 3.38; Score 44.6; DB 1; Length 36200;

Best Local Similarity 65.08; Pred. No. 3.6;

Matches 65; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 77 ttacgaagcgcgcgttcgtcggaagaataatgccgcatgccccctngatgaagca 136

Db 4337 TTTCGCACATCGGGCTGATACGCTGACATTTGACGCGACGCCCTTTCTTGAAGGCTRA 4278

QY 137 aggaagcgatgcgcgtcaaaaaagccaagtgcgtctgtga 176

Db 4277 AAGAAGCGATACCGCTCAAAAAAGCGAAGTGTGTTGA 4238

RESULT 22

BSUB0021/c 215534 bp DNA linear BCT 26-NOV-1997

DEFINITION Bacillus subtilis complete genome (section 21 of 21): from 399281

to 421481.

ACCESSION 299124 AL009126

VERSION 299124.1 GI:2636442

KEYWORDS

SOURCE Bacillus subtilis.

ORGANISM

Bacillus subtilis

REFERENCE

AUTHORS

Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus.
Bacillus/staphylococcus group; Bacillus.
1 (bases 1 to 215534)

AUTHORS

Kunst, F., Ogasawara, M., Moszer, I., Albertini, A.M., Alloni, G.,
Azevedo, V., Bertolo, M.G., Bessières, P., Bolojin, A., Borcherdt, S.,
Boriss, R., Boursier, L., Brans, A., Braun, M., Britnell, S.C.,
Bron, S., Brouillet, S., Brusch, C.V., Caldwell, B., Capuano, V.,
Carter, N.M., Choi, S.K., Codani, J.J., Conerton, I.F., Cummings, N.J.,
Dantel, R.A., Denicot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D.,
Emmerson, P.T., Entlian, K.D., Errington, J., Fabel, C., Ferrari, E.,
Fouger, D., Fritz, C., Fujita, Y., Fuma, S., Galizzi, A.,
Galleron, N., Ghim, S.Y., Glaser, P., Goffeau, A., Goldthly, E.J.,
Grandi, G., Guisepi, G., Guy, B.J., Haga, K., Haeich, J., Harwood, C.R.,
Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.F.,
Itaya, M., Jones, L., Joris, B., Katamata, D., Kasahara, Y.,
Klaer, Blanchard, M., Klein, C., Kobayashi, Y., Koelter, P.,
Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A.,
Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Liu, H.,
Masuda, S., Manuel, C., Medigue, C., Medina, N., Mellado, R.P.,
Mizuno, M., Mostl, D., Nakai, S., Nohack, M., Noone, D., O'Reilly, M.,
Ogawa, K., Ogihara, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M.,
Portetle, D., Portolillo, S., Prescott, A.M., Prescan, E., Pujic, P.,
Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M.,
Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadale, Y., Sato, T.,
Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J.,
Sekowska, A., Seror, S.J., Serron, P., Shin, B.S., Soldo, B.,
Sorokin, A., Taccioni, E., Takagi, T., Takahashi, H., Takemaru, K.,
Takeduchi, M., Tamakoshi, A., Tanaka, T., Terstra, P., Tognoni, A.,
Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vassarotti, A.,
Viari, A., Wambitt, R., Wedler, E., Wedler, H., Weitenegger, T.,
Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K.,
Yoshida, K., Yoshikawa, H.F., Zumbstein, E., Yoshikawa, H. and
Danchin, A.

The complete genome sequence of the gram-positive bacterium
Bacillus subtilis
Nature 390 (6657), 249-256 (1997)
98044033
2 (bases 1 to 215534)

Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.
Direct Submission
Submitted de 18-NOV-1997) I. Moszer, A. Danchin, Institut Pasteur,
Regulation de l'Expression Génétique, 28 rue du Docteur Roux, 75724
Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr,
adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45
68 89 48

FEATURES

source

Location/Qualifiers
1..215534
/organism="Bacillus subtilis"

gene

/db_xref="taxon:1423"

terminator

/db_xref="taxon:1423"

CDS

/db_xref="taxon:1423"

gene

/db_xref="taxon:1423"

terminator

/db_xref="taxon:1423"

CDS

/db_xref="taxon:1423"

gene

/db_xref="taxon:1423"

terminator

/db_xref="taxon:1423"

CDS

/db_xref="taxon:1423"

gene

/db_xref="taxon:1423"

terminator

/db_xref="taxon:1423"

CDS

/db_xref="taxon:1423"

gene

/db_xref="taxon:1423"

terminator

/db_xref="taxon:1423"

CDS

/db_xref="taxon:1423"

gene

/db_xref="taxon:1423"

terminator

/db_xref="taxon:1423"

TITLE Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1
JOURNAL Science 286 (5444), 1571-1577 (1999)
MEDLINE 20036896
REFERENCE 2 (bases 1 to 11375)
AUTHORS White, O., Eisen, J.A., Heidelberg, J.F., Hickey, E.K., Peterson, J.D., Dodson, R.J., Haft, D.H., Gwinn, M.L., Nelson, W.C., Richardson, D.L., Moffat, K.S., Qin, H., Jiang, L., Pamphile, N., Crosby, M., Shen, M., Vamathevan, J.J., Lam, P., McDonald, L., Utterback, T., Zaleski, C., Makarova, K.S., Aravind, L., Daly, M.J., Minton, K.W., Fleischmann, R.D., Ketchum, K.A., Nelson, K.E., Salzberg, S., Smith, H.O., Venter, J.C. and Fraser, C.M.
TITLE Direct Submission
JOURNAL Submitted (08-Nov-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source 1..11375
/organism="Deinococcus radiodurans"
/strain="R1"
/db_xref="taxon:1299"
/chromosome="1"
/gene="DRI1571"
complement(101..1885)
/gene="DRI1571"
complement(101..1885)
/note="similar to GB:AB000657 percent identity: 51.51; identified by sequence similarity: putative"
/codon_start=1
/transl_table=11
/product="peptide ABC transporter, periplasmic peptide-binding protein"
/protein_id="AAFI1135.1"
/db_xref="GI:6459337"
/translation="MKKVMALALAGSTSLAPFVYPANWTSNKPQDVOTGCTFRSY NLODEFTLNPVSESPNLPVLSAGSLGNTPTGVAAPYAKETYSADKRTTFPD IRKMKWSGDKPTTVDWMTIATYITDSKDKGSNTFDWTJINNOPIKATKVSNTLKVY FRKADYTALEFLSGITFLPOPTHFMPWKAKGAGIKDMWTISTPNDI VSGPPMLD RVRGKRAILKKNPYFGEMNKDSAGKSLPYLDGIQINIVADANQLOFLKGNLDIYS PKNRKLAKVKSAMDGKVKGTLPNARSASDFEYFNMDSATFKTSLSNVFRQ AMSMLNRDAMDALGIGLEPTYSVPYKDWI PSMGDKYKRPNTAAKLAELGF TKKSGDGLIVDAKNKLEFLITNAENRROSAYKVIODEAKKGVK INVASIARNO TLLDAKNRGRRNFDAITIGLGGGQVYSGSVCEKIGDGGNTLHMNSQKCR PPEFTQAVNLFWKGRAPFLAKRKATIAQIQRNEMENOPYIQLAQIVHFRMTDVRVG ENRPQINSLSMTLFGPRDIALTWIKR"
/gene="DRI1572"
/note="this region contains an authentic frame shift and is not the result of a sequencing artifact; similar to GB:AI009126 percent identity: 48.68; identified by sequence similarity; putative; helicase-related protein, authentic frameshift"
4529..6163
/gene="DRI1573"
4529..6163
/gene="DRI1573"
/note="similar to GB:M22039 SP:P13242 PID:143597 PID:853762 GB:AI009126 percent identity: 73.03; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="cnp synthase"
/protein_id="AAFI1136.1"
/db_xref="GI:6459338"
/translation="MKYIFVTGGVSVSLGKVASASIGALLRARGYKVTAVKIDPYIN IDACTMRPEHGECEFTASGAETDLDIGNERFLDIPAGSNITTGQVGEVIRKER AGDYLSQYOVVLPHTVDKIRIRAGESAGAEVLVEGTVGADIESIPLAEILROF RPEGGDDVYLHLITVPLTGSNERKRTQHSVALRSGISPDIVYVASKDKLPA ETRKIAFTSVRENRESESYDEYEHYOVPLAEQGLKAVEELDLEGTHPRLGVA TDAVETIKHPKOQVITAIAGKYTAMPDAVLISMESTHAGIANDARVNIKVNABOLA BAGAELEGOAFDVGILVGGFGIGIEGKVAEAERTHGYVLGICLGMQJAVIE YARHVAGIEDANSAEDEYAKNKVIDLMEOLEVAGSGTMRDMPMEIARAGTIAE LYGVGGGTGVEKRHRREYVNPAYVBOLEAGLTISGVTGCVBGRGALVETVELADH PEFVALQAHPEKSRMRSPPRAGVKALEHQOC"

gene 6206..6922
/gene="DRI1574"
6206..6922
/gene="DRI1574"
/note="similar to GB:U77117 SP:O58227 PID:1591506 percent identity: 49.48; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAFI1137.1"
/db_xref="GI:6459339"
/translation="MISASGQNVGRIARAYAPMRPLRRFPPLIAALAAAGMYLAVRY ERDVRPLPADOSGGAVALSPADGLVSEYVRRIEGGVNGONVAEILGTPRODGLGL YVGLDYHETAYAPYGGTVISAGKRCGTGOGPIATFARLAAQOPVLLGSAAVRGRS KSYAVGEGGQTVVGLIAPGAGLOAITYLDEGQTVRQGNKAFLAEGGLVLLALPES VTPQVSGERVGCAETVAC"
complement(6927..7574)
/gene="DRI1575"
complement(6927..7574)
/gene="DRI1575"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAFI1141.1"
/db_xref="GI:6459343"
/translation="MPRSELTITFOPLAIEHYDPCATLQTPGANIILNAAFPAPAE AENLPVYRWFSAKGRPALISDALIGEYDGLRVGSWTAGETRENTIVEVSRIQ LISMSAVLAESYETLEMAPALARHFAELGDRASTLLAYADGGVAGLWMTGGG TGAHLWGCTLDPADAPLNAADLSGQNVRSIDPSTPSVPEAEVLTFLDGS"
complement(7614..7937)
/gene="DRI1576"
complement(7614..7937)
/gene="DRI1576"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAFI1142.1"
/db_xref="GI:6459344"
/translation="MMFRDPPDITLPHAPLMQISQAQAYLECPAARDRAALL TALEDDYGAAPAPLAVTKLTGEAMQRTLPASEQTAAALLADFRAFLRGWMDARA VNLDP"
complement(7984..9072)
/gene="DRI1577"
complement(7984..9072)
/gene="DRI1577"
/note="similar to GB:M37702 SP:P21516 PID:147485 GB:U00096 PID:1773089 percent identity: 62.61; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="s-adenosylmethionine:trna rlpoyl transferase-isomerase"
/protein_id="AAFI1138.1"
/db_xref="GI:6459340"
/translation="MPDSTADPSTADEVLARLHPDLPEDRTIAOTGAEPDRTSRIMVYG EGIKRNFRDLPDLRLPSDLLENSESTIPARVARKVYNGFEGGQVEVLLIREED VGANVMSAYILKPAKRAGHEILGNEBRTGHAERYGVYLDGKRLRLRFHDITRPHIDEI GRLPLPYINAGDSDETWREKYQTVYATTPSSVAPVGLDGLTPPELLALDPMGVERA SVTLHVGAGTFKPIQGPVADHVMAERYEVSETNAATTRAKABGRVVAVGTTVVRT LESAMGGAARAGAGETRIFTTPTGRVREVPDLITNLHLPGSTLLLVAAFGEDRIR AAVDALISOYRFSLGDAMLERLTSA"
complement(9076..9633)
/gene="DRI1578"
complement(9076..9633)
/gene="DRI1578"
/note="similar to PID:1001613 PID:1001641 percent identity: 56.10; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11

AC005268 AC004330 AC004314 AC003599
AC005268.1 GI:3293207
HTG.

Drosophila melanogaster (Subclones in sac from pl clones DS03550 (D225) and DS02397 (D206)) DNA.

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 89791)
Celiker,S.E., George,R.A., Galle,R.F., Hoskins,R.A., Svirskas,R.R., Harris,N.L., Agbayani,A., Arcains,T.T., Baxter,E., Blazer,R.G., Chavez,C., Chew,M., Doyle,C.M., Fafan,D.E., Flanagan,J., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,S.H., Lee,B., Lomotan,M.A., Mak,J., Mazda,P., Mok,M.S., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacle,J.M., Park,S., Pfeiffer,B., Punch,E., Snir,E., Twomey,B., Wan,K.H., Whitehaw,K.R., Yee,A., Zhang,R., Zieran,L.L. and Kimmel,B.

Sequencing of *Drosophila* chromosome, region 57B6-57C1
Unpublished (1997)
2 (bases 1 to 89791)
Celiker,S.E., George,R.A., Galle,R.F., Hoskins,R.A., Svirskas,R.R., Harris,N.L., Agbayani,A., Arcains,T.T., Baxter,E., Blazer,R.G., Chavez,C., Chew,M., Doyle,C.M., Fafan,D.E., Flanagan,J., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,S.H., Lee,B., Lomotan,M.A., Mak,J., Mazda,P., Mok,M.S., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacle,J.M., Park,S., Pfeiffer,B., Punch,E., Snir,E., Twomey,B., Wan,K.H., Whitehaw,K.R., Yee,A., Zhang,R., Zieran,L.L. and Kimmel,B.E.

Direct Submission
Submitted (07-JUL-1998) Berkeley *Drosophila* Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US
Sequence submitted by:
Berkeley *Drosophila* Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720

For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://fruitfly.berkeley.edu/sequence/>) or send email to drosophila@hgsc.lbl.gov.
Library locations: 137-37, 136-25.
Location/Qualifiers
1. 89791
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="2R"
/map="57B6-57C1"
/clone="Pis DS03550 (D225) and DS02397 (D206)."
/note="DS03550 (d225) extends from bp 1 to bp 16,924 and DS02397 (D206) extends from bp 13,480 to bp 89,791."
24290 a 20016 c 20807 g 24678 t

BASE COUNT
ORIGIN

Query Match 2.8%; Score 38.2; DB 3; Length 89791;
Best Local Similarity 46.4%; Pred. No. 1.3e+02;
Matches 121; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

DB 1072 aaaaacaaactctcaagttcaagcaagcgctcaagcgtgtgacgacgcgcgcattgtgcg 1131
DB 78595 AACCAACACATACGCAAGTCATCCATCCATCGTGTCTATGCGCACATCAAACTTGAAGCAG 78654
QY 1132 atgtgacttaagagcgcgtaaagcgcttagacacatcccgctaacctgtcttgcgcat 1191
DB 78655 GATCTGAAGGCATCATGTCGACAGCATCTCGAGCGATGAACCAACATGCCCACATCCGCGTG 78714
QY 1192 ttaatcgttcgcgatataccagcagcgcgcaagcattgggtgtcgttgtaattgacgaaga 1251
DB 78715 CTTCAAGATTGGGGGTCCGAGGTACGCTATTCGCCCGGTCAACTACGAAGGAACCTTAAC 78774
QY 1252 gacctcgcttctgacgctctgctcccgggcaaatagaatlangccgcgtgtgtgt 1311

Db	78775	GCCGGGAGGAGGACGCACCGCTGCTATTCCAGAGTCTCTCCTTCGATGTGCGGGCGAAGCCGG	78834
Qy	1312	aagtgctggaacacnttgag	1332
Db	78835	CAGAGACTGCAATCCCTTGCGC	78855
RESULT	30		
AC018341/C		AC018341	101085 bp
LOCUS		Drosophila melanogaster, ***	DNA linear HTG 09-DEC-1999
DEFINITION		pieces	
ACCESSION		AC018341	
VERSION		AC018341.1	GI:6552850
KEYWORDS		HTG; HGS; PHASE2.	
SOURCE		fruit fly.	
ORGANISM		Drosophila melanogaster	
REFERENCE		Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
AUTHORS		1 (bases 1 to 101085)	
TITLE		Adams, M. and Venter, J.C.	
JOURNAL		Direct Submission	
COMMENT		Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA	
FEATURES		This sequence was identified as CDM:10214413 by the submitter. For more information on this record e-mail to fly@celera.com.	
source		* NOTE: This is a 'working draft' sequence.	
		* This sequence will be replaced	
		* by the finished sequence as soon as it is available and	
		* the accession number will be preserved.	
		Location/Qualifiers	
		1..101085	
		/organism="Drosophila melanogaster"	
		/db_xref="taxon:7227"	
BASE COUNT		28299 a 22687 c 22278 g 27821 t	
ORIGIN			
Query Match		2.8%; Score 38.2; DB 2; Length 101085;	
Best Local Similarity		46.4%; Pred. No. 1.3e+02;	
Matches 121; Conservative		0; Mismatches 140; Indels 0; Gaps 0;	
Qy	1072	aaaacaacattcaattcacagacagcagcgtacacggtgagcagccagtggtccg	1131
Db	92552	AACAACACATACGCAAGTCCATCCATCGTGTCTATGSCCACTTACACTTTAAGCAG	92523
Qy	1132	atgtgacttaacagagcgcgtlaatgcgcgtacacalcctgcctacccctgtcttgcgcgat	1191
Db	92522	GATCGAAGGCGATCGATCGACAGCAATCGTGGACGCTGATACCAATCGCAATCGCCATCTGGTG	92463
Qy	1192	ttaatcgtcgcgagataaccagcagcgcgcgaagcattgggttctgttgaaattgacgaagaa	1251
Db	92462	CTTCAGGATGTGGCGGCGAGTACCGCCATATCCGCGGCTCAACTACGAGGAACCTTAAC	92403
Qy	1252	gaccctcgctgttgagcttgcctgcgcgcgcgcaaatgaaatlangccgcgctgttgct	1311
Db	92402	GCCGGCGAGGAGCGACCGTGTATTCCAGAGTCTGTCTTCGATGTGCGGGAACGGCG	92343
Qy	1312	aagtgctggaacacnttgag	1332
Db	92342	CAGGAGCTGAATCCCTTGCGC	92222
RESULT	31		
AF263912		AF263912	123580 bp
LOCUS		Streptomyces noursei ATCC 11455	DNA linear BCT 24-MAY-2000
DEFINITION		Streptomyces noursei ATCC 11455	nystatin biosynthetic gene cluster,
ACCESSION		AF263912	
VERSION		AF263912.1	GI:8050835
KEYWORDS		Streptomyces noursei.	

ORGANISM	Streptomyces noursei
REFERENCE	Bacteria: Firmicutes: Actinobacteria: Actinobacteridae: Actinomycetales: Streptomycineae: Streptomycetaceae: Streptomyces. 1 (bases 1 to 123580)
AUTHORS	Brattaset,T., Sekurova,O.N., Sletta,H., Ellingsen,T.E., Strm,A.R., Valla,S. and Zotchev,S.B.
TITLE	Biosynthesis of the polyene antifungal antibiotic nystatin in Streptomyces noursei ATCC 11455: analysis of the gene cluster and deduction of the biosynthetic pathway
JOURNAL	Chem. Biol. 7 (6), 395-403 (2000)
MEDLINE	20334850
REFERENCE	2 (bases 1 to 123580)
AUTHORS	Brattaset,T., Sekurova,O.N., Sletta,H., Ellingsen,T.E., Strom,A.R., Valla,S. and Zotchev,S.B.
TITLE	Direct Submission
JOURNAL	Submitted (04-MAY-2000) Unigen, NTNU, O. Kyrras gt. 3, Trondheim N-7489, Norway
FEATURES	Location/Qualifiers
source	1..123580
gene	/organism="Streptomyces noursei"
CDS	/strain="ATCC 11455"
gene	/db_xref="ATCC:11455"
CDS	/db_xref="taxon:1971"
gene	complement(46..783)
CDS	/gene="nysF"
gene	complement(46..783)
CDS	/gene="nysF"
gene	/function="presumably post-translationally modifies ACP domains on PKS"
CDS	/note="putative 4'-phosphopantetheine transferase"
gene	/codon_start=1
CDS	/transl_table=11
gene	/product="NysF"
CDS	/protein_id="AA71762.1"
gene	/db_xref="GI:8050836"
CDS	/translation="MTELIPATVTEAYDDRRPRGDRLLSEREVYIARVESRQREFTVRIARLARLRIGHDPRIALPNRGAPOWPGIVGSMTHCAGYRAAASPAELSAVSIDAPNGPLPAGVLAIALPSEPHILALAHARDVMDRLLEFSAKESYFAWYPLQRELDSEAEIVIDPTOGAFARLLPGLGRRVTVFPGRMHSTPALLTAVHLPAPIPRDRHRHLTVNSPLPRPTFG"
gene	complement(867..2684)
CDS	/gene="nysG"
gene	complement(867..2684)
CDS	/gene="nysG"
gene	/function="presumably involved in efflux of nystatin"
CDS	/note="putative transporter (ABC family)"
gene	/codon_start=1
CDS	/transl_table=11
gene	/product="NysG"
CDS	/protein_id="AA71763.1"
gene	/db_xref="GI:8050837"
CDS	/translation="MASPDLEERTAPRPARRVGLLRPRRSVALAVSMGVGIVLNAFGRLLGRVMTLLDGVGVRGAPRGIDFAALIGRLILVLLVYASLPMACQRLVASAVWRTIHELRLDAREKULRLPHRQRPAGELLSRTINDINLOQLQOTLALNHSIFSLTMLVLMLEISPLAVSLAVESALIAARISKRAQPHAAQWMSANGLNHAEVCTGHALIKGFDRRAAEERDANDAVYRAAKQFASGAMPEMVMVANGLYLVAVVIGAMVINGNTLIGVOAFLIYAROFISPIVEIASVAGRLQSGJASAOVFTLLDAPEOAPDLRPGTPARAREGVEFTDVSFRYSPDPLEIENSLTYEESYTAIVGPGCACTTIGNLMLRVEYRPSGRILLDGTATATRRDIRSRGCIYADPTLFGCTIAENIATGACGACRADIEBARATCADRFIRTLPGQDVLVDDSEGTVSAGEQDLTVARALAPRAVILVDEATSSVDTRTEVLIGRAMNSLNAGRSPVIAHRLSLRDADLIVMDAGRIEQTGHDLDLCAOGLYARLHAARTHTPTAGAAG"
gene	complement(2662..4416)
CDS	/gene="nysH"
gene	complement(2662..4416)
CDS	/gene="nysH"
gene	/function="presumably involved in efflux of nystatin"
CDS	/note="putative transporter (ABC family)"
gene	/codon_start=1
CDS	/transl_table=11
gene	/product="NysH"
CDS	/protein_id="AA71764.1"
gene	/db_xref="GI:8050838"

gene	/translation="MLRLRLAQLRPYAWATAALVALDQVILGTLLEPLTGAALIDQGVNRDGRITFELGVYMAVYVQIAALAAALATATATAGRDLSALFRILIDSAREIGRFGPPLSLFTRVNDVNOVONIAOGFIVCAPIAMCGSVLLALRODPLLSLVALYVAVCGILLARMGTLKARMOQLLNLGGLRALTIGVVRFSVDHERRA RFAQNDNFELVSVRVGRILATMLVYLLMLANGFTYALLMTGSHRIDARMPIGSIALLSYLSILMSVYMAFVFLSVPRARVCGRIAEVYALDDTSVAAPAPDPVKPARI ELCAAGYRPGAEFEPLRVLDLVEBEREIAVIGSGSCKTLLNLVYLAATEAVR RGGVDVEIRFATATLAAGVFPYORPVEFSYVSNLRFGRPDATDELMALATVQA ADVFARMPDGLDAEITOGGNSGGORSLARLLRREPEITLPDSCSALDOATDA ALRTALVPYTAGATYTVAGRISACRDADRIYVLDGRVYAGCTHVLRTSTYREI ALSQLEEREAHGLAGR"
gene	4714..5748
CDS	/gene="nysDIII"
gene	4714..5748
CDS	/gene="nysDIII"
gene	/function="presumably involved in mycosamine biosynthesis"
CDS	/note="putative dGDP-mannose-4,6-dehydratase"
gene	/codon_start=1
CDS	/transl_table=11
gene	/product="NysDIII"
CDS	/protein_id="AA71765.1"
gene	/db_xref="GI:8050839"
CDS	/translation="MSKRALLTIGTGQDGSYLAHLLISGQYQWGLIRGOANRKSRY SRLASELDFIDGLNDQGSLSVAVDTPQDEVTYNLALISFVPMKQOAEVTVYNGKGLRLMEALRMVSGLSSTSRVSPRGQIRFYQASSSEMFGAETPQRETTLFRSPYGLAKAAGYHITRNYRESFGMYAVSGMLFNHESPRGGEFTYRKISLVARIKQGLQDL AAGNLDAVRDMGYAGDYVRAMHLMLOODQGDVDVITGQMHSVDAVARIFEHGLMM EDVYVIDPOLVPAEVEYLCADSAKAKQDBLGMKPDVDPTLMAMVDSDLAQSRENO YGDVLIANM"
gene	5930..34363
CDS	/gene="nysI"
gene	5930..34363
CDS	/gene="nysI"
gene	/function="responsible for condensation steps 9 to 14 in the nystatin polyketide backbone synthesis"
CDS	/note="polyketide synthase"
gene	/codon_start=1
CDS	/transl_table=11
gene	/product="NysI"
CDS	/protein_id="AA71766.1"
gene	/db_xref="GI:8050840"
CDS	/translation="MDNEOKLRDYKLATRADLRTRRVYKLESAAQEPALITGTCR YPGVRSPELDMKMEVAGEHGVTFPTDKMDLEALAAATFASGTLHDPDADPAG GISPRAVAMPDPOQRYVLESAMAEFRAGIDPWSVSGRTGVFTGAADAVRGPAG AEGFOLTGNTGTVLSGRISYTFGTVPAYVDVACSSILVAAHIALALAGECTLL AGGVTMSRGTFIEMRGQGSADCRSGFDATGTMAEGLILVLRSDAVAN GHEILLAVRGTAVNODGANGLTPARPGSQOQYIOQALVVARLAAGIDVVEAGHGT TIGDPEVQALATYGCNRRADRPPLLGYSKSNLSHTQAAAGYATKMYAMHGTLP RLIAEPEPTHNVDSGAVRLRLTDITDMFATGAPRAAASFTGISNTNHTTIEORP EPOPEDAATADDAAGSTPATAPVAVGVVLLSGRPDALRQQAALRAALDGRRP DLIDLALHSLATRACGEHRAVILLATDHPALTDLGTLADDDPAAPAMVGTTRARP RLAVLETTGOGAOLVGLRELAAREPAFATVLTDLADLAFTHDRLPDLVEVLTGDAALL DRTAAYOAPALFNEVALYRLIESEGRPHLADHSGEIVIAHLAAGLSLADATTLVA ANGRLQALPDGAMATVAQASEADVAPLLAGHDOVAIAVNGPNSVYLSGAEVTYVA LAEQILDARCKTRLVNVAHFSPLEMPMLDARVAVEDLIDPILPVP SNLTGKRA TYAQLLSADYVNDVHRAVAFADGIDMLADHDTAFLEGPDLVLSMADDCDAAATA DAVTLEGTADGPREERTLTTLAAGLVHGAFTDMGTGCFAGTGARRDLPYAFGRRY WKXALDSGADRSVGLGAHNPBLNAAVSLADAGCTILGRLSROTHNPMYLDITVGG TLLPBTGAFLLELVARAGDEYGCGRVEETLALDLLPEQGVQVOLLIGNDVSGRPT VYVNHARPDTGDDTPWTAHATGVTTTADASQOLPASSSQGCTPLAGDPHPLDAAQMPR ACAPERLPUGHNDRLDAGGFGTVPVQGLAARNGGVYVAEBELPAGSDBAEFLI HALIDALAAHAEFTGILGERGRLPFSWMSGJLHAGATLTLRLPVDADLALTY ADGTAPVLSVDSLTYSVATQDLDTAAVARALPFLMDTPVQPTATDGPVALYLA DEFGTLHAGFADAPAYPDLAALAADGPVPTTVVLSLCTGDADPARSANGCLAE ALAAVQOTWLDHERRFAARLVFVTRGATVGRDVAAAAMVGLVRSAGSENGCFALYLD DPGAGCAALVYALYSGERQLAVRDVLEVAALVRRPLTEVAGAGCTGDGVDGSG VSPSGEAVLVYGTGGLGAVLRLHAEYGVNDLLVYSGEAVACAGELVATLACY GARVRVACDVYDRAAVVELVGHAAVAVVHAGVLDGAVGALTGRLSALVLPKYD AWWHLEATRGIDLDAFVYSSLAGVSGGQANVYAAANAFILAAMTRRRECEYGLS LAMGPWEOSSGGMGTLLDVAERLASVGSPLVSAQALFEADAVMTDADTACVYVRL DPLTRAGGEVPPILRLIRYARRAAVAGSATGNLAKQRLRLDEDEGRDEVLDLVVG OVALYIGHATGCDVYDGRAPRDIGFBSLTVVELRNRLNTYTGRLPRTVTFDYTVRH LATYVLDLDELGTDAEVAIVOPAAVADDPVIVAGMACRYPGVSSPBDMLRVLTBGT

OY	1192	ttatctgcggcgagataccgacgaagcgcgaagcatgttggttctgttgaaattgagcaaga	1251
Db	203031	CTTCAGGATGGCGGGGTCCGAGTACGCCCTATTCCCGCGGCAACTACGAACGAACCTAACCC	202972
OY	1252	gacctgcgatttgttcagacttcgcttcgcccgaggacaatacgaataingccscgctgtgcgt	1311
Db	202971	GCCGGCAGGAGACGCACCCTGCTATTCCAGAAGTCTGCTCTTGATGTGCGGCGAAGCGGG	202912
OY	1312	aaggtgcctggaacacnttgaq	1332
Db	202911	CAGGAGCTGAAATCCCTTGCGC	202891
RESULT	35		
AC110159		69283 bp DNA linear HTG 10-FEB-2002	
DEFINITION	MUS MUSCULUS clone RP23-4P22, LOW-PASS SEQUENCE SAMPLING.		
ACCESSION	AC110159		
VERSION	AC110159.1 GI:18643511		
KEYWORDS	HTG; HTGS; PHASEO.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
JOURNAL	Biren,B., Linton,L., Nusbaum,C. and Lander,E.		
TITLE	1 (bases 1 to 69283)		
REFERENCE	Mus musculus, clone RP23-4P22		
AUTHORS	unpublished		
	2 (bases 1 to 69283)		
	Biren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,		
	Anderson,S., Barina,N., Bastien,P., Boguslavsky,L., Boukhalter,B.,		
	Brown,A., Camarata,J., Campotroiano,A., Chang,J., Chazaro,B.,		
	Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,		
	Cooke,P., Dextrallano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,		
	Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,		
	Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,		
	Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,		
	Kamat,A., Karatas,A., Kells,C., Lakocue,K., Lamarez,R.,		
	Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,		
	Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,		
	McDonald,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,		
	Mlinga,V., Murphy,T., Naylor,T., Nguyen,C., Nicoll,R., Norbu,C.,		
	Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,		
	Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,		
	Retta,R., Ribback,M., Riley,R., Rise,C., Rogov,P., Roman,A.,		
	Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnoback,R., Seaman,S.,		
	Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,		
	Strauss,N., Subramanian,A., Talamas,J., Testafay,S., Theodore,J.,		
	Totham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,		
	Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.D., Young,G.,		
	Zainoun,J., Zemek,L., Zimmer,Zody,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-FEB-2002) Whitehead Institute/MIT Center for Genome		
COMMENT	Research, 320 Charles Street, Cambridge, MA 02141, USA		
	All repeats were identified using RepeatMasker:		
	Smt, A.F.A. & Green, P. (1996-1997)		
	http://ftp.genome.washington.edu/RM/RepeatMasker.html		
	Genome Center		
	Center: Whitehead Institute/ MIT Center for Genome Research		
	Center code: MIBR		
	Web site: http://www-seq.wi.mit.edu		
	Contact: sequence_submissions@genome.wi.mit.edu		
	Project Information		
	Center project name: I17557		
	Center clone name: 4_P_22		
	*		
	* NOTE: This record contains 89 individual		
	* sequencing reads that have not been assembled into		
	* contigs. Runs of N are used to separate the reads		
	* and the order in which they appear is completely		
	* arbitrary. Low-pass sequence sampling is useful for		
	* identifying clones that may be gene-rich and allows		

DVALALTRDLAIVSLTRASDALYLHELEDGSLRAAGLSAFLDAGALAELEKVEPADIDR
VVAEQAPEPLPADGIPEADQVPPFCPTRELEEVAGAGHNYHADLINVTEGEREVR
YMRISRLHKNHTEMPTGEREVLAVCAVRRYRAGEDESTLTLAARONPREPROPP
PRVTAGVAOEMRTMYLERIDLTIVYOMGKGLERIDYRVARREYVSTPESFCAHTOIS
VPAFLKATLKCVDAALGPRDEPCCHAAOGKRALEIRAKAEVVOYMSHFRMKTOKIM
RALRPPFLVAAGHTEREVDAMWQHHTTNALVEBTEEDMOTLTATRVDELISALL
GLPCADYERALARAGSYCTLRELSTEGCESTSEPAFLIHTNTVAMCMAMMVPAGV
RMAGIFQGDVMVIFLPEGARSALAKMTPAEVGLGFHPIVKHVSSTPESFCGHVIAA
GLEHDVMOAIKIVLCREFDPVLEEOVALLDRIKGVYAAALPDVMAANAAYDYSAER
VLAIVRELTAIVANGRGIDHPATIGALEIEIOTPYARANLHDDAD"
6493..9684

CDS

/codon_start=1
/product="structural proteins C, E2 and E1"
/protein_id="AAL31568.1"
/db_xref="GI:16923735"
/translation="MASTRPITMEDLOKALEAOSRALRAELIAGAOSRRPPRRPO
SSTSGDSDSGSGRRRRRNGRGRKDMAMPREPERESQOTAPAKSRAPPO
OPQPPMOTGRGSSAPRELGPPTNPFOAAVARLRLPLDPTDTEAPTEACTYSLWS
EBGAVFYRYDLHFTNLGTPRLDEDGMDPALMPCGEPERAHVRAVNOFAGDVRG
VWGKGERTYEODFRVGGTRMHLRMPVRGLDSDAPLRPHETERIETRSARHWPRI
RGAPQAFLAGLILAAVAVGTARAGLQPRADMAAPAPROPCHAGOHYHHNOLPF
LGHDHGGTLRVGOHHRNASDYLPGHMLGGMGCCYNLSDMHOGHTVHTKMDMVCY
EHDRPPATPTLTAA NSTTAATPATAPACCHAGLNDSCGFLSGCCPMRLRGADT
RCGRILCGISTTQIYPTPRACAMKCLPMEILVLLARPEDGWCRCGVPAHPTGCP
ELVSPMRATCSPASALWLTANALSLDHLAALVLEPWLIFMVCRRACRRGAAA
ALTAVALQYINPAGYEATYLCCTARCATQAPVPAVLACVRESKIVDGCFAAMD
LEATGACICEIPTDVSCIEGLAWPTAPCARIMGTOACTFEWNAYNSGGYADLAS
YNPGGSYKYOYATACEVPAFGHSDACMGEPDYMVSFALASYOHHPKAYRVK
FHETRTVMOLSYAVGSCNTPTEHPRCNTHPGOLFVOYPPRPGDVEIIMNTDQOS
RMGLGSPNCHGPRMSPVPCORHSDCRSLVGAITERPRLVDDADPLRTAAGVEV
WTPYVIGSARCKGLHTRAGPYGHAVEMPEMNAHRTSDPMHPCPGIGLRFKTVAPV
ALPRALAPRMAVNTCYOCGTALVEGLAPGGCNHCLTVNGEDVGAAPRGKFTVVAL
LMTPTQVYSCGSEDSRASARVLDPAOSFTGVYVGHHTTAVSETROTWAEWAANAHHW
OITLGAVALLAGLILACCAKCLYLRNGATAPR"

BASE COUNT 1448 a 3761 c 3007 g 1527 t
ORIGIN

Query Match 2.8%; Score 37.2; DB 14; Length 9743;
Best Local Similarity 45.7%; Pred. No. 2.2e+02;
Matches 123; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 55 gtcattatgaagcgccgcgtcattaccgaagtcgcgttgcgttcggaagaatbaccggt 114
Db 3935 GGCACCTGTGCCCGGACGCGAGGGGCTAGCCGAGCGGTACTACGACGACCTCGAGGTG 3984
QY 115 atgcgccccctnagatgaagaagcgaagcgaatgcctcaaaaaagcgaagtcgtgtt 174
Db 3985 CGCCGCTCGGGGATACGCCCATAGGCCCGGCGCCCTCGCATGATCTCAACGCCCTCGC 4044
QY 175 gaagacaaaaagatccggcgctgtgttaccgcgcgcngtttcaggaacaaatcgccgc 234
Db 4045 AAAAGCCCCCTACAAATATACGAGGTATGGAACATGGCCGACGCGCGGCAAGCACCCGC 4104
QY 235 atccatcgagggaagcgaagcgaatcagtcggtcgtgtatccgcgttgaagcgaagc 294
Db 4105 ATCCCTGCTGCTTCACGGCGGCAAGACCTCTACGCTTGCCTCCCAATGCGCTCCTGCAC 4164
QY 295 gaatcgaatcgaagcgaagcgaagcga 323
Db 4165 GAGATCCAGCCCAAACTTGGCGGCGCGA 4193

RESULT 38
AF435866 9743 bp RNA linear VRL 14-NOV-2001
LOCUS AF435866 Rubella virus isolate Surkova, complete genome.
ACCESSION AF35866
VERSION AF35866.1 GI:16923736
KEYWORDS
SOURCE Rubella virus.
ORGANISM Rubella virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;

REFERENCE
AUTHORS 1 (bases 1 to 9743)
TITLE Rubella virus (Isolate 'Surkova') genomic RNA
JOURNAL 2 (bases 1 to 9743)
UNPUBLISHED
REFERENCE
AUTHORS Hofmann, J., Renz, M., Meyer, S., von Haeseler, A. and Liebert, U.G.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-2001) Institute of Virology, University of
Leipzig, Max-Buehner-Forschungszentrum, Johannsallee 30, Leipzig
04103, Germany

FEATURES

source

CDS

/codon_start=1
/product="nonstructural proteins NSP1 and NSP2"
/protein_id="AAL31567.1"
/db_xref="GI:16923737"

CDS

/codon_start=1
/product="structural proteins C, E2 and E1"
/protein_id="AAL31568.1"
/db_xref="GI:16923738"
/translation="MASTRPITMEDLOKALEAOSRALRAELIAGAOSRRPPRRPO
SSTSGDSDSGSGRRRRRNGRGRKDMAMPREPERESQOTAPAKSRAPPO
OPQPPMOTGRGSSAPRELGPPTNPFOAAVARLRLPLDPTDTEAPTEACTYSLWS
EBGAVFYRYDLHFTNLGTPRLDEDGMDPALMPCGEPERAHVRAVNOFAGDVRG
VWGKGERTYEODFRVGGTRMHLRMPVRGLDSDAPLRPHETERIETRSARHWPRI
RGAPQAFLAGLILAAVAVGTARAGLQPRADMAAPAPROPCHAGOHYHHNOLPF
LGHDHGGTLRVGOHHRNASDYLPGHMLGGMGCCYNLSDMHOGHTVHTKMDMVCY
EHDRPPATPTLTAA NSTTAATPATAPACCHAGLNDSCGFLSGCCPMRLRGADT
RCGRILCGISTTQIYPTPRACAMKCLPMEILVLLARPEDGWCRCGVPAHPTGCP
ELVSPMRATCSPASALWLTANALSLDHLAALVLEPWLIFMVCRRACRRGAAA

RESULT 43
AC027658/c 142737 bp DNA linear PLN 02-APR-2000
LOCUS Oryza sativa subsp. japonica BAC nbxb0006113, chromosome 10,
DEFINITION complete sequence.
ACCESSION AC027658
VERSION AC027658.1 GI:7363409
KEYWORDS HTG.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eurharoidae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 142737)
AUTHORS Llaça,V., Young,S., Kavchok,S., Clark,T., Choi,V.,
Nevill-Manning,C. and Messing,J.
TITLE Rice Chromosome 10
JOURNAL Unpublished
2 (bases 1 to 142737)
AUTHORS Llaça,V., Lou,A., Young,S., Charydzac,G., Choi,V.,
Nevill-Manning,C. and Messing,J.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2000) The Plant Genome Initiative at Rutgers -
Wakeman Institute, Rutgers University, 190 Frelinghuysen Road,
Piscataway, NJ 08873
3 (bases 1 to 142737)
AUTHORS Llaça,V., Young,S., Kavchok,S., Clark,T., Charydzac,G., Choi,V.,
Nevill-Manning,C. and Messing,J.
TITLE Direct Submission
JOURNAL Submitted (02-APR-2000) The Plant Genome Initiative at Rutgers -
Wakeman Institute, Rutgers University, 190 Frelinghuysen Road,
Piscataway, NJ 08873
REMARK Location/Qualifiers
FEATURES
source 1..142737
/organism="Oryza sativa"
/db_xref="taxon:4530"
/chromosome="10"
/clone="nbxb0006113"
BASE COUNT 42049 a 28618 c 30051 g 42019 t
ORIGIN
Query Match 2.8%; Score 37.2; DB 8; Length 142737;
Best Local Similarity 47.5%; Pred. No. 2.3e+02;
Matches 105; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
Oy 107 atgcccgtatgccccttngatgaagtaaggaaggaagcgatgcgcgtcaaaaagccaaag 166
Db 42381 ACCTCTCTGCTCGGCGACGTGTGCTCAAGATCGCGACGTCTCCCGCCGACGGCG 42322
Oy 167 tgcctgttgaagcaaaaagatccggcgctggtgttaacgcgcgcngtttaagcaaaa 226
Db 42321 TGTTCCTGGACGGCGCACGGCCTCGAGTCGACGAGTCGACGATGACCGCGACGGCAC 42262
Oy 227 tcgcgcgcacatccatcgcgcggaagcgctacttcagtcgctgtgattgcgctgaag 286
Db 42261 CCGTCGAGGTGCGACGGCGTCGTCAGAGCCCTTCCTCGGCTCGGGGTCAGAGTGCACG 42202
Oy 287 gcaacgaagaatcgagttcgaaagcgtacgcgcgcgaagcg 327
Db 42201 GCTACGGCAGATGTGTACCGCGCTCGGAGCGACGACG 42161
RESULT 44
NMA52491 329861 bp DNA linear BCT 04-DEC-2000
LOCUS NMA52491
DEFINITION Neisseria meningitidis serogroup A strain Z2491 complete genome;
segment 5/7.
ACCESSION AL162756 AL157959
VERSION AL162756.2 GI:7380091
KEYWORDS

SOURCE
ORGANISM Neisseria meningitidis Z2491.
Neisseria meningitidis Z2491
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria
REFERENCE 1 (bases 1 to 329861)
AUTHORS Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
Davies,R.M., Davis,P., Devlin,K., Feltham,T., Hamlin,N.,
Holroyd,S., Jagsels,K., Leach,S., Moule,S., Mungall,K.,
Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
Skellon,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
TITLE Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491
JOURNAL Nature 404 (6777), 502-506 (2000)
MEDLINE 20222556
REFERENCE 2 (bases 1 to 329861)
AUTHORS Parkhill,J.
TITLE Direct Submission
JOURNAL Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENTS Details of N. meningitidis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES
source 1..329861
/organism="Neisseria meningitidis Z2491"
/strain="Z2491"
/db_xref="taxon:122587"
/note="serogroup: A"
194..565
/gene="NMA1439"
194..565
/gene="NMA1439"
/note="NMA1439, possible lipoprotein, len: 123 aa;
unknown, contains a probable N-terminal signal sequence
and an appropriately positioned PS00013 Prokaryotic
membrane lipoprotein lipid attachment site"
/codon_start=1
/transl_table=11
/product="putative lipoprotein"
/protein_id="CAB84675.1"
/db_xref="GI:7380092"
/db_xref="SPTRMBL:Q9JU92"
/translation="MNKTLSTLPVALILGGCAAGNTFGSLDGGNGMSYKMAVE
SOCRAELNKRESEMTLALMSAEKAEENNICACVAQAPNQLGNDVMQMLDPSTR
NOALATARTVSCFKHLIR"
212..244
/gene="NMA1439"
/note="PS00013 Prokaryotic membrane lipoprotein lipid
attachment site"
complement(571..580)
/note="Core DNA uptake sequence: gccgctctgaa"
/label-DUS
592..601
/note="Core DNA uptake sequence: gccgctctgaa"
/label-DUS
638..2047
/gene="thrc"
638..2047
/gene="thrc"
/EC_number="4.2.99.2"
/note="NMA1440, thrc, probable threonine synthase, len:
485 aa; similar to many e.g. SW:THRC_METGL (EMBL:D14071),
thrc, Methylobacillus glycoenes threonine synthase (EC
4.2.99.2) (475 aa), fasta scores: E(): 0, 61.9% identity
in 475 aa overlap. Contains pfam match to entry PF00291
S.T dehydratase, Pyridoxal-phosphate dependent enzymes"
/codon_start=1
/transl_table=11
/product="putative threonine synthase"
/protein_id="CAB84676.1"

DEFINITION	Sequence 13 from Patent WO9940109.
ACCESSION	AX019436
VERSION	AX019436.1 GI:10043375
KEYWORDS	
SOURCE	
ORGANISM	synthetic construct. artificial construct. artificial sequence.
REFERENCE	1 (bases 1 to 763)
AUTHORS	Welters,P., Stiefel,T., Voelter,W. and Morris,P.
TITLE	Recombinant mistletoe lectines
JOURNAL	Patent: WO 9940109-A 13 12-AUG-1999; WELTERS PETER (DE); STIEFEL THOMAS (DE); BIOSYN ARZNEIMITTEL GMBH (DE); VOELTER WOLFGANG (DE); MORRIS PETER (GB) Location/Qualifiers
FEATURES	1..763
source	/organism="synthetic construct" /db_xref="taxon:32630"
misc_feature	319
BASE COUNT	163 a 215 c 175 g 170 t 40 others
ORIGIN	

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2002, 07:14:05 ; Search time 27321.5 Seconds
(without alignments)
1029.420 Million cell updates/sec

Title: US-09-303-518D-131
1344
Sequence: 1 atgatacaatacaaaaagg.....ccattgagaagaagcctga 1344

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 100 summaries

Database : GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_ov: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htgo_inv: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1344	100.0	1344	6	A96098	A96098 Sequence 13
2	1212.8	90.2	1344	6	A96092	A96092 Sequence 12
3	1212.8	90.2	12273	6	AE002412	AE002412 Neisseria
4	1212.8	90.2	349880	6	AX044029	AX044029 Sequence
5	1179.2	87.7	311321	1	NMA322491	AL162754 Neisseria
6	1165.4	86.7	1344	6	A96094	A96094 Sequence 12
7	574.4	42.7	10563	1	AE006171	AE006171 Pasteurell
8	566.2	42.1	1558	1	APU24492	U24492 Actinobacill
9	550.2	40.9	13299	1	U32702	U32702 Haemophilus
10	518.4	38.6	5916	1	AF117331	AF117331 Vibrio ch
11	518.4	38.6	13382	1	AE004300	AE004300 Vibrio ch
12	513.6	38.2	6048	1	AF165980	AF165980 Vibrio ha
13	488	36.3	5078	1	VANQRBOL	237111 V.alginolyt
14	483.2	35.0	6908	1	AB008030	AB008030 Vibrio al
15	475.6	35.4	15903	1	AE004724	AE004724 Pseudomon
16	433.6	32.3	474	6	A96090	A96090 Sequence 12
17	348.8	26.0	96109	6	AX067460	AX067460 Sequence
18	126.4	9.4	1356	1	AF144076	AF144076 Porphyrom
19	105.4	7.8	4926	1	H1020229	U20229 Haemophilus
20	42.4	3.2	36200	1	BACGNTZA	D78193 Bacillus su
21	42.4	3.2	151329	2	AP003456	AP003456 Oryza sat
22	42.4	3.2	215534	1	BS0B0021	299124 Bacillus su
23	40.2	3.0	125020	9	AF429315	AF429315 Homo sapi
24	40	3.0	37000	14	BHT10L	Z78205 Bovine hept
25	40	3.0	135301	14	BHT10L	AJ004801 Bovine he
26	39.8	3.0	216050	1	AL646076	AL646076 Ralstonia
27	39	2.9	125020	9	AF429315	AF429315 Homo sapi
28	38.8	2.9	8918	1	AB047272	AB047272 Pseudomon
29	38.8	2.9	123580	1	AF263912	AF263912 Streptomy
30	38.8	2.9	125401	6	AX211739	AX211739 Sequence
31	38.4	2.9	10500	1	AE005062	AE005062 Halobacte
32	38.2	2.8	3000	1	AF220510	AF220510 Azarcus
33	38.2	2.8	9980	1	AE004915	AE004915 Pseudomon
34	38.2	2.8	38734	1	SC10A5	AL021529 Streptomy
35	38	2.8	1641	1	AB015974	AB015974 Pseudomon
36	37.8	2.8	340900	1	SME591791	AL591791 Sinorhizo
37	37.6	2.8	760	4	AY011741	AY011741 Canis fam
38	37.2	2.8	1731	8	KP081143	U81143 Klebsiella
39	37.2	2.8	4130	8	AF190427	AF190427 Colletotr
40	37.2	2.8	7218	6	166494	166494 Sequence 14
41	37.2	2.8	213050	1	AL646079	AL646079 Ralstonia
42	37	2.8	10029	1	AE008180	AE008180 Agrobacte
43	37	2.8	10359	1	AE009214	AE009214 Agrobacte
44	36.8	2.7	10514	1	AE002070	AE002070 Deinococc
45	36.4	2.7	3855	13	M23555	M23555 Figure 2. N
46	36.4	2.7	3867	1	AY033407	AY033407 Myxococcu
47	36.4	2.7	189050	1	AL646069	AL646069 Ralstonia
48	36.2	2.7	30132	1	SC10G8	AL158057 Streptomy
49	36.2	2.7	110000	2	LMFLCHR32_04	Continuation (3 of
50	36.2	2.7	174301	2	AP003523	AP003523 Oryza sat
51	36	2.7	11526	1	AF228498	AF228498 Escherich
52	35.8	2.7	16897	1	AE006941	AE006941 Mycobacte
53	35.8	2.7	35019	1	MTY13E10	295324 Mycobacteri
54	35.8	2.7	154330	2	AC104429	AC104429 Oryza sat
55	35.8	2.7	157912	9	AL133387	AL133387 Human DNA
56	35.6	2.6	632	3	AF240033	AF240033 Asterias
57	35.6	2.6	960	1	PSU83882	U83882 Pseudomonas
58	35.6	2.6	1462	1	PSRCATEX	L10655 Sphingomonas
59	35.6	2.6	1918	1	AB008137	AB008137 Enterobac
60	35.6	2.6	1930	1	AB008140	AB008140 Enterobac
61	35.6	2.6	4004	1	SSP284817	284817 Sphingomona
62	35.6	2.6	12129	1	AE007223	AE007223 Sinorhizo
63	35.6	2.6	17205	1	AE007185	AE007185 Mycobacte
64	35.6	2.6	24268	1	AB070947	AB070947 Streptomy
65	35.6	2.6	30352	1	MTCY409	297188 Mycobacteri
66	35.6	2.6	41321	1	MSG4409	AD000017 Mycobacte
67	35.4	2.6	296	6	AX260544	AX260544 Sequence
68	35.4	2.6	10732	6	E32986	E32986 Gene encodi
69	35.2	2.6	3955	14	NFOEGT	D17353 Oryzia pseu
70	35.2	2.6	10234	1	AE009378	AE009378 Agrobacte
71	35.2	2.6	10647	14	AE008237	AE008237 Agrobacte
72	35.2	2.6	131995	14	OP075930	U75930 Oryzia pseu
73	35.2	2.6	152941	8	AP003044	AP003044 Oryza sat

74	35.2	2.6	162010	2	AP004121	AP004121 Oryza sat
75	35	2.6	388	12	AF316400	AF316400 Synthetic
76	35	2.6	1261	9	AF316401	AF316401 Homo sapi
77	35	2.6	1563	9	AF316402	AF316402 Homo sapi
78	35	2.6	1631	9	BC012134	BC012134 Homo sapi
79	35	2.6	1634	9	BC006413	BC006413 Homo sapi
80	35	2.6	1700	9	AF043250	AF043250 Homo sapi
81	35	2.6	1724	9	BC001779	BC001779 Homo sapi
82	35	2.6	1731	9	BC017224	BC017224 Homo sapi
83	35	2.6	7870	9	HS275213	HS275213 Homo sapi
84	35	2.6	11374	1	AE009451	AE009451 Brucella
85	35	2.6	11501	1	AE009276	AE009276 Agrobacte
86	35	2.6	12674	1	AE008336	AE008336 Agrobacte
87	35	2.6	42655	1	SC7H2	AL109732 Streptomy
88	35	2.6	189050	1	AL646066	AL646066 Ralstonia
89	35	2.6	213861	9	AL135793	AL135793 Human DNA
90	34.8	2.6	355	11	G72966	G72966 MARC 2567-2
91	34.8	2.6	975	1	AF103943	AF103943 Streptomy
92	34.8	2.6	1325	6	AX168124	AX168124 Sequence
93	34.8	2.6	3105	1	HALHMA	M87833 Halobacteri
94	34.8	2.6	10689	1	AE000430	AE000430 Escherich
95	34.8	2.6	14000	6	AX353826	AX353826 Sequence
96	34.8	2.6	19088	1	AE001334	AE001334 Chlamydia
97	34.8	2.6	225419	1	ECOW76	U00039 E. coli chr
98	34.6	2.6	1780	8	TAAMY1	X98504 T.aestivum
99	34.6	2.6	2968	1	MSG6081	U6081 Mycobacteri
100	34.6	2.6	16796	1	AE002269	AE002269 Chlamydia

ALIGNMENTS

RESULT 1
LOCUS A96098
DEFINITION Sequence 131 from Patent WO9924578.
ACCESSION A96098
VERSION A96098.1 GI:6779941

SOURCE
ORGANISM unidentified.
unclassified.
unclassified.
REFERENCE 1 (bases 1 to 1344)
AUTHORS Pizsa,M., Scarlato,V., Rappuoli,R., Grandi,G. and Masiugnani,V.
TITLE Neisserial antigens
JOURNAL Patent: WO 9924578-A 131 20-MAY-1999;
PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);
CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT);

FEATURES
source Location/Qualifiers

1..1344
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 334 a 363 c 362 g 285 t

Query Match 100.0%; Score 1344; DB 6; Length 1344;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgattaaatcaaaaaggtctaaatcgcgcattcgcggagcagccgagcgaatcatt 60
Db 1 ATGATTAAATCAAAAAAGGTCTAAATCTGCCATCGCGGACGACCGGAGCAAGTATT 60
QY 61 tatagcgcgcgcgcattaccgaagtcgcgttcgttcgcgaagaatatgtcgcgcatgc 120
Db 61 TATGACGGCCGCGCATTAACCAAGTCCGCTTGCTTGCGGAGCAATATGTGCGCATCGC 120
QY 121 cctcgatgaataatcaaggaagtgaaagccgtcaaaaaaagcgaagtgcgtttgaagac 180
Db 121 CCTCGATGAATAATCAAGGAAGTGAAGCGCTCAAAAAAGGCAAGTGTGTTGAAGAC 180
QY 181 aaaaagaatccggcgctgatttactcgcgcgcgttcgaagaataatcgcgcattcac 240

|||||
Db 181 AAAAAGATCCGGCGTAGTATTACTCGCCGGCTTACGGAATAATGCCGTATTAC 240
QY 241 cgtgagaaagcgcgtcattcagtcagtcgtgattgcgttgaaggaacgaagaatc 300
Db 241 CGTGGCCAAAAGCCGCTACCTTCAGTCACTCGATCGGTGCGTTGAAGGCAACGCAAAATC 300
QY 301 gattcgaagctacgtacccggaagcgttcgcaaatgtgagaaggaagaagtgcgcgc 360
Db 301 GAGTTCAGACGCTACGTACCTGAAGCGCTGCAAAATTGAGCAGCAAAAGTGCCGCC 360
QY 361 aacctgattcaatcaagccttatgactgcgttcgcacccgcgttcgaacaaatccct 420
Db 361 AACCTGATTCAATCAGGCTTATGAGCTGCGTTCGACCCGTCGTTACGAAAAATCCCT 420
QY 421 ggcgtatgcgcgcgcgttcgcgcattcgttcgaatgcgatgagacaacaatcgtgcct 480
Db 421 GCGGTGATGCCGAGCCGCTTCGCATCTTCGTAATGATGAGACCAATCCGCTGGCT 480
QY 481 ggcgacctacgtcattcatcaagaagcgcgcgaagcttcaacgcgcgttcgttgta 540
Db 481 GCCGACCTTACGGTCAATCATTAAGAAACCGCCGAGACCTTCAAAACGGCCTGTGGTA 540
QY 541 ttgaagcgcctgaacgcgaacgtlaaaatccatgtgtgtaagcagcagcgcagcgtgcgc 600
Db 541 TTGAGCCGCTGACCGCAAGTAAATCCATGTGTGTTAAGCAGCAGCGCGACAGCTGCCG 600
QY 601 tctgaaatgtctccaatatcgaacaacatgaatttgcgcgcgcgcgcgcgcgcgcgcgc 660
Db 601 TCTGAANAATCTGCCAATATCGAANAACATGAATTTGGCGGCCCATCTCCGCTTGG 660
QY 661 agtgcacgcacattcatcattcatcgaacgaatcgcgcgcgaataaaacgcgttgaccatc 720
Db 661 AGTGCACGCAATTCATTTTCATGAGCCATCTCGCGGATTAACCGTGTGAGCAATC 720
QY 721 aattatcaagcgttgattgtctatcgaacgttgcgttcgtaacagcgcgttcgaataccgag 780
Db 721 AATTATCAAGACGTGATTCATCGACGTTTGTCTGAACAGCCGCTGTGAATACCGAG 780
QY 781 cgcgtgttccttcggcgcgccgcaagtaacaacacgcgccttcgttcgacgtttg 840
Db 781 CGCGTGTTCCTTGGCGCGCTGCAAGTCAACAACCGCCCTCTTGCATCCGTTTGG 840
QY 841 ggtgcgaaggtgtctcaactaccgcgcgcgaattgtgtgcgcgcgaacacgcgtgatt 900
Db 841 GGTGCGAAGGTGTCTCAACTTACCGCGGGAATTGTTGACGGGACCAACCGCGTGATT 900
QY 901 tccggttcggtattgaacggttcgatttcgaacaagcgcgcgcatgatttctggagcgtac 960
Db 901 TCCGTTTCGTTATGAACGCTGCGATTGCACAAGCGCGCATGATTATTGAGACGCTAC 960
QY 961 ccaaatcagaattccgttatcgaagaagcgcgcgaagaagcgttcgttcggttcg 1020
Db 961 CACAAATCAGATTTCCGTTATCGAAGAGCGCGCACAGCAAAAGCTGTGTGGCTGGGTTGG 1020
QY 1021 ccgcagccgcgaacaatactcatcaccgcgcacacactcgcgcattcctctaaaaaaa 1080
Db 1021 CCGACCCGACAAATATCTCATTCACGCGCACCACTCTCGGCCATTCTTAAAAAACAA 1080
QY 1081 cctcttaagtctcgaacgcgcgtcaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1140
Db 1081 CTCTTCAAAATTCAAGACAGACGCGTCAAGCGGCGACCGCGCCATGTATCCATCGGCACT 1140
QY 1141 tatgaacgcgtatgcgttcgttcgaacatcctcctaaccttcgttcgttcggaattta 1200
Db 1141 TATGACCGGCTAATGCGCTTGAGACATCTCTTACTTGTTCGGGATTAATTCGTC 1200
QY 1201 ggcgaataccgaacgcgcgcgcgcgttcgttcgttcgttcgttcgttcgttcgttcgttc 1260
Db 1201 GCGGATACCGACAGCGCGCAGGCTTGGTGTGGAATTGGAGCAAGAAAGACCTGCT 1260
QY 1261 ttgtgcagcttcgttcgcgcgcgcgaataacgaatacgcgcgcgcgcgcgcgcgcgcgcgc 1320

Db 1261 TTGTGACGCTTGTGCGCGGCAATAAGCAATACGGCCGCTGTGCGCAAGTGTG 1320
Qy 1321 gaaccatttgagaagaagctga 1344
Db 1321 GAACCACTTGAAGAAGAGCTGA 1344

RESULT 2
A96092 1344 bp DNA linear PAT 07-SEP-2000
LOCUS DEFINITION Sequence 125 from Patent WO924578.
ACCESSION A96092
VERSION A96092.1 GI:679938
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1344)
AUTHORS Pizza, M., Scarlato, V., Rappuoli, R., Grandi, G. and Masiugnani, V.
TITLE Neisseria antigens
JOURNAL Patent: WO 924578-A 125 20-MAY-1999;
PIZZA MARINERAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);
CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)
FEATURES
source 1.1344
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 329 a 366 c 365 g 284 t
ORIGIN

Query Match 90.2%; Score 1212.8; DB 6; Length 1344;
Best Local Similarity 93.9%; Pred. No. 3.2e-295;
Matches 1262; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 1 atgattaaatcaaaaaaggctcaatctgccaatcgcgagcaagcgaagcgaatcatt 60
Db 1 ATGATTAAATCAAAAAAGGCTCAACCTGCCATCGGGGAGACGCGACCAACCCGTT 60
Qy 61 tatgaagcgccgagcattacacgaagtcgctgttggaagaataatgctgcacgcgc 120
Db 61 TACGAGCGCGCGGCGCATACGGAAGTCGCGTGTGGCGAAGATATGCGGATGCGC 120
Qy 121 ccttcgatgaataatcaagaagtggaagcgttcaaaaaagcgaagtcgtttgaagac 180
Db 121 CCTTCGATGAAGTCAAGGAAGGATGCGCTCAAAAAAGGCCAAGTCTGTTTAAAGAC 180
Qy 181 aaaaagaatccggcgtagtaattactgacgagcgttcgaagcaaaatcgccgtatac 240
Db 181 AAAAAGATCCGGCGTGTGATTACGCGCGGCTTCAGGCAAAATCGCGCGATTAC 240
Qy 241 cgtgagcaaaaagcgctactcaagtcagtcgttgatgctgttgaagcaagcaaatc 300
Db 241 CGTGGCGAAAAAGCGCTTCACTCACTGATTCGCTTGAAGGCAACGCAAAATC 300
Qy 301 gaattgaaagctacgctacgaagcgttgcaaaattgagcagcgaagaaatgctgcgc 360
Db 301 GAGTTTGAAGCTACGCACTGAAGCGCTGCAAACTTAAGCGGCGAAGAACTGCGCC 360
Qy 361 aaactcattcaatcgaactatgactgagtcgcttcgacccgttcgcttcagcaaaa 420
Db 361 AACCTGATCAAAATCCGCTTGTGACGCGCTGCGACCCGTCGTTACGCAAAATTC 420
Qy 421 gcgtagatgcgagcgctgcgcacattcgtcaatgcgaatggaacaaatcgcgtgct 480
Db 421 GCGGTGATCCCGAGCGCTTGCATCTTCATATGGAATGGAACCAATCCGCTGCT 480
Qy 481 ggcgaacctacgctcatcaaaagaagcgcgcaagactcaaaagcgcgctgttgta 540
Db 481 GCGGACCTTACGCTCATTTATCAAAAGAACCCCGAGATTTCAAAACGCGCGCTTGGTA 540
Qy 541 ttgagcgctgacccaagcaatcaatcgaatgltgttaagcaagcagcgacagctgcgc 600
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Db 541 TTGAGCCGTTTGAACGGAACGAAATCATGTTTGAAGGACAGTGGCGACAGTGGCC 600
Qy 601 tctgaataatgctgcgaataatcgaacacatgaatttggcgccgcgcatccttcgcgctg 660
Db 601 TCTGAATAATGCTGCACAACTTCGAACACATGAATTTGGCGCGCCGATCTCTGCGGTTTG 660
Qy 661 agtgcagcgaatcattcattcagcagtcgagtcgagcaataaaacgctgtgagcattc 720
Db 661 ACTGGCAGCCACATTCATTCATTCAGAGCGGCTGCGCGCAATTAACCGTGTGACCATC 720
Qy 721 aattcaagaagctgattgctatcgagctgttgcgttaacagcgcttcgaataccgag 780
Db 721 AATTATCAAGATGATTAATTCATTCATTCAGTGGCGCTGTTTGGCAACAGCGCTGGAAC 780
Qy 781 cgcgtgattgctgctgagcgccgcgaagtcaaaaacccgcctcttgcgtaccgtttg 840
Db 781 CCGGTGATTGCGCTTAGGTGTTTCAAGTCAACAAACCGCGCTCTTGGCTACCGTTTG 840
Qy 841 ggtgcgaaggtgtctcaacttaccgcgagcaattgtgtgacgcggaacacgcgctgatt 900
Db 841 GGTGCAAAAGTATCCGAATTAATCTCGGCGGGAATTTGTTGACACAGAACCGCGTAT 900
Qy 901 tccggttcgctatgaagctgctgattgcaagcgcgcgatgattatgttgagcgttac 960
Db 901 TCCGTTCCGTTATGAACGCGCGGATTAACAAGCGCGCACATTAATTTGGGACGCTAC 960
Qy 961 caaatcagatttcgcttaccgaagaagcgcgcaagaagcgaagcgttcgctgctgctg 1020
Db 961 CAAATCAGATTTCCGTTATTCGAAGAAGCGCGCACGAAGAAGCTGTCGCTGCGTTGCG 1020
Qy 1021 ccgagcgcggaacaatactcattcagcgcgcaacactctcgccatttcctaaacaaa 1080
Db 1021 CCGCAGCGCGCAAAATACATCATCACGCTGTAACAACCTCGGCGCTCTGTAACAA 1080
Qy 1081 cctcttaagttcagcaagcgctcaagcgcgagcgcgacgctgctgctgctgctgct 1140
Db 1081 CTCTTAAGTTCAACACAGCGCTCAACGCGCGCACCGCGCATGCTGCGATTGTA 1140
Qy 1141 tatgagcgctgaatgcgcttggaacatcctcactcctgcttcttgcgagatlaactgct 1200
Db 1141 TACGAGCGCTGATGCCCTTGGATATCTGCGCACCGCTGTTTGGCGGATTAATGCT 1200
Qy 1201 ggcgaatcagcagcgcgagcgttggctgttgccttgaattggaagcaagacccgct 1260
Db 1201 GCGGATACGACAGCGCGGCGCATTTGGTGTGGAATGAGCAAGAAAGACCTCGCT 1260
Qy 1261 ttgtcagcttcgctgcgcccgggaataatgaataagcgccgctgttgcaaatgtgtg 1320
Db 1261 TTGTGAGGCTTGTGCTGCCGCGGCAAAATACGAATACGCGCGCTGTGCGAAAGTGTG 1320
Qy 1321 gaacccattgagaagaagctga 1344
Db 1321 GAACCACTTGAAGAAGAGCTGA 1344

RESULT 3
AE002412/c 12273 bp DNA linear BCT 25-MAY-2000
LOCUS DEFINITION Neisseria meningitidis serogroup B strain MC58 section 54 of 206 of the complete genome.
ACCESSION AE002412 AE002098
VERSION AE002412.1 GI:7225783
KEYWORDS
SOURCE Neisseria meningitidis MC58.
ORGANISM Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 12273)
Tetelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C., Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F., Dodson, R.J., Nelson, W.C., Gwinn, M.L., Deboy, R., Peterson, J.D., Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D., Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Cittone, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H.,
Qin, H., Vanathavan, J., Gill, J., Scarlato, V., Masignani, V.,
Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappuoli, R. and Venter, J.C.
Complete genome sequence of *Neisseria meningitidis* serogroup B
strain MC58
Science 287 (5459), 1809-1815 (2000)
2 (bases 1 to 12273)
10710307
Tettelin, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,
Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
Dougherty, B.A., Mason, T., Clecko, A., Parksey, D.S., Blair, E.,
Cittone, H., Clark, E.B., Cotton, M.D., Utterback, T.R., Khouri, H.,
Qin, H., Vanathavan, J., Gill, J., Scarlato, V., Masignani, V.,
Piazza, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
Rappuoli, R. and Venter, J.C.
Direct Submission
Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
1. 12273
/organism="Neisseria meningitidis MC58"
/strain="MC58"
/db_xref="taxon:122586"
/note="serogroup: B"
97. .435
/gene="NMB0557"
97. .435
/note="NMB0557"
/note="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF40985.1"
/db_xref="GI:7225784"
/translation="MSDEPILFTDSCCAKVAADLAENNNPDKLRVFNAGGCGSEFO
YGFPEDEIKNDDEFEIKRGLVLPDPMSTYLVGAKEIDYTESLQSGSYIRPNMET
TCGCGSSPSV"
595. .1059
/gene="NMB0558"
595. .1059
/note="NMB0558"
/note="hypothetical protein; identified by Glimmer2;
putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAF40986.1"
/db_xref="GI:7225785"
/translation="MIEMSKDYRNDLYDVVYSPPOVDRLIRECLKENLGEKARGLT
IESLSKPOVLEVEKETWAKRELDHYSYGLDIITRRYMELEFVYPPREBEGBEG
ADGMPETELTEHGREDDISAPSOPEPSRNKILVFGILIAFLGTLLEKIP"
1094. .2605
/gene="NMB0559"
1094. .2605
/note="NMB0559"
/note="similar to GB:M87049 SP:P27853 PID:148233 GB:U00096
PID:2367309 percent identity: 74.73; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="ubiquinone biosynthesis protein AarP"
/protein_id="AAF40987.1"
/db_xref="GI:7225786"
/translation="MKMLKRLVIVGFYRRYRLAGLCASLWGSNICALIKMMPQSSK
LKNPPAVRLALASLGIPIKGOVISTRPDLIPHYAVELAKIDKVPPEPDLAS
REQIETKSIGSIEKLYAEFEPEPIASASIAOVHKLRLSGEQAVALRPNLIPVTEO
DLISMRGAGVERLPADGKRLPREVVAEFDKYLDHDLIREAANAOLCRNONS
DMLVPRVFYDYCTSDVLTIRMMDGTPVSDIAKLKADGIDLKLADYGEVEIFTOVFR

gene
CDS
DGEFHADHPCGIIIVAADNRVIALDQGVIGTLFDYDKRYLAINEFLAFNRDVRVATA
HIESGVADPRAEELAEAVRACVEPRFKRPSQISGLVNLAEVSRFENEIQPO
LVLLQKTLNIEGLDELDLMTKATPVLKWMNGGVGRALNRNLKNAAPDAQ
IIPSLPRKISALIDENROEMRDAYIHWKVOORSLMVAIVALLITLLK"
complement(2962. .3780)
/gene="NMB0560"
complement(2962. .3780)
/note="NMB0560"
/note="similar to GB:LA2023 SP:P43886 PID:1005429
PID:1220683 PID:1204857 percent identity: 72.97;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="serine acetyltransferase"
/protein_id="AAF40988.1"
/db_xref="GI:7225787"
/translation="MKDHLNTGFDLMTHTREETAAYSAEPMASFLHQTVLRHES
LGSVALVHSSKSLSPIMDVRLAFPEIYQOALGSDIOIGCVADDAIYERPADCEY
SLPLIYEFKFAIOAHRIINHRSLYEDGRATIAFYLRNMSVEGADIIHPARAGIML
DHATGEVAGETAVIGNNISILHGYVLGSGKGGDRHPKIGGVWIGANASTIGNRI
GSMKIGGSVVSVPSTIVGVPAKPAVARSKLTSPSDMDQNIQFTIDPMT"
3964. .4542
/gene="NMB0561"
3964. .4542
/note="NMB0561"
/note="similar to GP:3777486 percent identity: 61.03;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="grpe protein"
/protein_id="AAF40989.1"
/db_xref="GI:7225788"
/translation="MSBOTQOQSEAVENYEAFTVETVGNAGDVORAAAPRYED
LQARIATFRLQKDEQLBALNEONLRRRHQDEIDTRKPAQOKRAVELPKYDLEM
ALDQSGNFDALKMGVQMTLNELOKAFDATOIKETNPAGDKLDPNIHQAMQAVASQ
EPNTVGVGMRKGYLSDVRLPAAVTAOKBA"
complement(4819. .5034)
/gene="NMB0562"
complement(4819. .5034)
/note="NMB0562"
/note="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF40990.1"
/db_xref="GI:7225789"
/translation="MKTLLTFGLFVLIYIGMAVGYIFSKRSKSGCGITALGKKM
CDCTPCDTLQKLDKROAGGIRVDR"
complement(5060. .6115)
/gene="NMB0563"
complement(5060. .6115)
/note="NMB0563"
/note="similar to GB:LA2023 SP:P44550 PID:1003244
PID:1222087 PID:1204428 percent identity: 74.70;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="thiamine biosynthesis lipoprotein AppB"
/protein_id="AAF40991.1"
/db_xref="GI:7225790"
/translation="MPSETRLPNFRILFALGFIFLNACSEQTAQVTLLOGEMGT
YTVYLSNNRDLKPSPAEIOKRIDDALKEVNRQSTOPEISRFNPHTKGKPLRIS
SDFAHVTAFAVRLNRLNLHGAIDVYGLVLMGSGPKSVYRREPSPDIKQAASTGT
DKTIILOKGRYASLSKTHPKAYLDLSIAKGFYDKVAGLEKYGIONYVEIGSELH
GKRNARGEFWRIGIEQPNIVQGGNTQIIVPLNRSIATSGDYRIFHYDKNGKRLSHI
INPNKRPIISHNLASISVVASMTADLSTGLVLTETELKLAERKLAIVFLIVRD
KGGYRTAMSSEPEKLIR"
complement(6268. .7485)
/gene="NMB0564"
complement(6268. .7485)
/note="NMB0564"
/note="similar to GB:LA2023 PID:1003242 PID:1222086

Query	Match	Best Local Similarity	93.9%	Score	1212.8	DB	1	Length	12273
Matches	1262	Conservative	0	Mismatches	82	Indels	0	Gaps	0
QY	1	atgataaatacaaaaagctctaatactgcgcacatcgcgagacacgaggaagtcatt	60						
Db	12069	ATGATTAAATCAAAAAAGCTCTAAACCTGGCCATCGCGGACGACGAGCAAGCCGTT	12010						
QY	61	tatgacggccggccattaccggaagtcgctgtgttggcgaagaatattgtcgatgcgc	120						
Db	12009	TACGACGGCCCGGCGCATTAACGAAATCCGTTGGCGGAATATCCGATGCGC	11950						
QY	121	ccctcgatgataaatacaaggaagtgtaagccgtcaaaaaagagcgaagtgcgttgtaagac	180						
Db	11949	CCCTCGATGAAGTCAAGCAAGCGCATCCGTCAAAAAGCGCAAGTGCTGTTGAAGAC	11890						
QY	181	aaaaaagatccggcgctagctattacatcgccgcgcttcgaagcaaatatgcgctattacc	240						
Db	11889	AAAAAAGATCCGGCGCTGCTGTACTGCGCGCGCTTCAGGCAAAATCCCGGATTCAC	11830						
QY	241	ctgtggcaaaaagcggtactactcaatcgatcgatgttcgcttgaaggaagcaagcaaatc	300						
Db	11829	CTGTGCGAANAAGCGGCTACTTACGTACGTACGTGATTTGCCGTTGAAGGCAACGAGAAATC	11770						
QY	301	gagtcgaacgcctagctacttgaaacgcgtgcaaaatgtagcagcagaaaaagtgcgcgc	360						
Db	11769	GAGTTTGAACGCTACGCACCTGAAGGCGCTGGCAAACTTAAGGGGCGCAAGAAAGTGGCCGCG	11710						
QY	361	aacctgattcaatcaggtctatgagctgagcttcggaaccgcttcgtagcaaaatccct	420						
Db	11709	AACCGATTCCAATCCGGTTTGAGATGGCGCTGGCACCCTGCTCCGTTACGCAAAATTCCT	11650						
QY	421	gccgtatagatgcgcgaacgcttcgcacatcttcgttaacatgcatgtaacacaaatccgctgct	480						

Db	11649	GCCGTCGATGCGAGGCCGTTCCGCAATTTGTCATGTCGATGACACCAATCCGCTGGCT	11590
Qy	481	ggcgacccttaagctcatcatcaagaagacgcgcgaagacttcaaacgcgcctgttga	540
Db	11589	CGCGACCCCTAGCGGTCAATATTCAAAGAACGCCGCCAGATTTCAAAACGGCGCTTGGTA	11530
Qy	541	ttgaagccgcctgaacccaagcttaaatccatgtgtgttaagaaagagcgagagctgcg	600
Db	11529	TTGAGCGCTTTGACCGGACCCAAAATCCATGTTTGTAAAGGACGTGGGCGACACGTGGC	11470
Qy	601	ttcgaanaatgctcccaatatacgaaacacatgaatttgcgcgcgcgccttgcgcgttgc	660
Db	11469	TTCTGAANAATCTGCGCAACATCGAAGAACATGAAATTCGGCGGGCGCCATCCTCGGTTTG	11410
Qy	661	agtgcacgcacatcatlctcatcgagcagtcgagcgcgatataaacogtvtgagacatc	720
Db	11409	AGTGCACGCAACATTCATTCATCGAGCGGTCCGCGGGAATAAACCGTGTGGACATC	11350
Qy	721	aattatcaagaacgtagtctgtatcaggaagtttgttcgttaacagcgctctgaataccag	780
Db	11349	AATTATCAAGTGTGAATTAATTCATTGCGCGTGTGTTGCAACAGCGCGTGTGAACACCGAG	11290
Qy	781	cgcgtgtgttccttgccttgcgcgcgcctgcgaagtcacaacacgcgcctcttgcataccgttgc	840
Db	11289	CGCGTGTATGCGCCTAGGAGGTCTTCATCAATCAACAAACCGCGCTCTTGGTACCGTTTG	11230
Qy	841	gtgtgcgaagtgcttcaactataccgcgcgcgaattgtgttgaacgcgcgaacacgcgtagt	900
Db	11229	GGTGCGAAGTATGCAAAATTAATCTGCGGCGAATTTGTTGACACAGACAAACCGGTGATT	11170
Qy	901	tccggttcgcgtatgaagcgtgcgtatgtcaacaagcgcgcatgatatttggagcgctac	960
Db	11169	TCGCGTTGCGTATGGAACGGCGGCGATTCACAAAGCGCGCACGATTAATTTGGACGCTTAC	11110
Qy	961	cacaatcaagatttccgttatatgaaagaagccgcgacgaagagctgttgcgtgtgttgcg	1020
Db	11109	CACAAATCAGATTTCCGTTATCGAAGAGCGCCGACGAAAGCTGTTCGCTGGATTGCG	11050
Qy	1021	ccgcgaacgcggaacaatctcatcacaacgcgcacacactctgcgcatttccataaaaaaa	1080
Db	11049	CCGCAAGCCGGAACAAATTAATCTCAATCAACGGGTACAAACCTCGGCATTTCGTAACAAA	10990
Qy	1081	ctcttcaagtccaagacagccgtcaacgcgcgcgcgcgcgcgtatgataccgactgcgac	1140
Db	10989	CTTTTCAAGTTTCACACAGCCCTCAACGGCGGCGACGCGCACTGTGTGCGGATTGGTACT	10930
Qy	1141	taabagcgcgttaatgcgcttggacatccctcactccttgccttcttgcgcgatttaactgic	1200
Db	10929	TACGAGCGCGTGATCCCTTGATATTCGTGCCACCTGTTTGGCGGATTAATGTC	10870
Qy	1201	ggcgataacgaagacgcgcgcgcgttggattgtcttgaatttggaaagaaagacgtcgtc	1260
Db	10869	GGCGTATACCGACAGCGCGGCGCATTTGGTGTGGAATTTGACGAAGAACACTGTGCT	10810
Qy	1261	ttgtgcagcttgcgtccgcgcgcgaataatcgaaatacgcgcgcgttgttgcgcaagtgctg	1320
Db	10809	TTGTGCACGCTTGCTGTGCCCGGGCAATATGAGATAGGGCCGCGTGTGGCAAGTGCTG	10750
Qy	1321	gaaacatttgaagaagaagctga	1344
Db	10749	GAACCAATTGAGAGGAAGGCTGA	10726
RESULT	4		
LOCUS	AX044029/c	349980 bp	DNA
DEFINITION	Sequence 108 from Patent W0006791.	linear	PAT 24-NOV-2000
ACCESSION	AX044029		
VERSION	AX044029.1	GI:11342913	
KEYWORDS			
SOURCE	Neisseria meningitidis.		
ORGANISM	Neisseria meningitidis.		
	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;		

REFERENCE 1 (bases 1 to 349980)
AUTHORS Piza, M., Hickey, E., Peterson, J., Tettelin, H., Venter, J.C., Masignani, V., Galeotti, C., Mora, M., Ratti, G., Scarselli, M., Scarlato, V., Rappelli, R., Frazer, C.M. and Grandi, G.
TITLE Neisseria genomic sequences and methods of their use
JOURNAL Patent: WO 0066791-A 108 09-NOV-2000;
FEATURES CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)
source location/Qualifiers
1. 349980
/organism="Neisseria meningitidis"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces, -seq 1: 1 to 349980 349980 bases-seq 108: 300001 to 649980 349980 bases-seq 109: 600001 to 949980 349980 bases-seq 110: 900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980 349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq 113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to 2272325 172325 bases"
BASE COUNT 84410 a 84863 c 94187 g 86520 t
ORIGIN

Query Match 90.2%; Score 1212.8; DB 6; Length 349980;
Best Local Similarity 93.9%; Pred. No. 4.8e-295;
Matches 1262; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 1 atgataataatcaaaaagtcataatctgcccacatcgccgagccagccgagccatcatt 60
Db 297909 ATGATTAATAATCAAAAAGCTCTAAACCTCCATCGCGGCGAGCCGAGCCGCTT 297850

QY 61 tatgagccgcgcgcacataccagagtcgcgcgtctctgccaagaataatgctgcgcatcgc 120
Db 297849 TACGAGGCGCGCGCCATTACCGAAGTCGCGTCTTGCGCAAAATATGCGGTTATGCGC 297790

QY 121 cccctgagtaaaatcaagaagtgagagccgctcaaaaagccgaagtgcgtttgaagac 180
Db 297789 CCCTCATGAAATCAAGAGGAGGCGATGCGTAAAAAAGCCAGCGCTTTGTAAGAC 297730

QY 181 aaaaagaatccgcgtagtattactgcccgcgcgttcacgaataatgcgcgtatcac 240
Db 297729 AAAAATAATCGCGCGCTGCGTGTACTGCCCGCTTCAGCGCAAAATGCGCGGATTAC 297670

QY 241 cgtgcgcaaaagcgcgtactcagtcagtcgagtcgctgcaagcaagcgaagcaatc 300
Db 297669 CGTGGCGAAAGCGCGTACTTCACTGATCGTGTATTCCTTGAAGCGCAAGCAATC 297610

QY 301 gagttgcaagcgcgtactcagtcagtcgcaaaaatlgagcagcgaaagtcgcgcgc 360
Db 297609 GAGTTGAAACGCTACCGACCTGAAGCGCTGCAAACTTAAGCGCGAAGAAATGCGCGC 297550

QY 361 aaactgataatcaagcgtatgagtcgctgcgacccgctcgttcagcaaatccct 420
Db 297549 AACCTGATCAATCCGCTGTTGTGAGTGCCTGCGCACCGCTCGTTACCAAAATTCCT 297490

QY 421 ggcgtgtagtcgcaagcgttcgacatcttcgtaatgagtcgagcaaatccgcgtgct 480
Db 297489 GCGGTGAGTCGAGCGCTTGCATCTTGTCAATGCGATGAGCACCAATCCGCTGCT 297430

QY 481 ggcgacccaatgctcatcaagaagccgcgaagactcaaacgcgcgtgtgtgta 540
Db 297429 GCGGACCCATACGCTATTTCAAAGAAGCGCGAGAGATTCAAGCGCGCTGTGTGTA 297370

QY 541 ttgagccgcgtcgaacgtaaatccatcgtgtgtaagaagcagcgagcgagcgtgcg 600
Db 297369 TTGAGCGCTTTGACCGCAACCAATCCATGTTGTAAGGACACTGCGCGAGAGCTGCCG 297310

QY 601 tctgaagaatgctccaatatcgaaacatgaattggcgccgcgacccctcgcgcgtgtg 660
Db 297309 TCTGAAATGCTCTCAACATCGAAACATGAATTCGCGCGCGCATCTCGCGGTTTG 297250

QY 661 agtgcgacgacatcattcatcgagcagtcgcgcgcaataaacgctgtgacatc 720
|||||

Db 297249 AGTGGACACGCACATTCATTTCATCGAGCGCGTCCGCGAATAAACCGTGTGACCATC 297190

QY 721 aattatcaacgctgattgctatcgagcgttctgctgaacagcgccgtgaatacagag 780
|||||

Db 297189 AATTATCAAGATTAATTACATTGGCGGCTTTGTTGCAACAGCGCGCTCGAACACCGAG 297130
|||||

QY 781 cgcgtgctgcttgaggcgccgctgcaagtaacaacaacgcgcctctgtgacgctttg 840
|||||

Db 297129 CGCGATTGCGCTAGGTGCTTCAAGTCACAAACCGCGCTCTGCTACCGTTTTG 297070
|||||

QY 841 ggtgcgaagtgctcaactaccgcgcgcggaattgctgaecgagcaacccgctgatt 900
Db 297069 GGTGCGAATATATCGCAATTAATCTGCGCGGGAATGTGTTACACAGAACACCGCTGAT 297010
|||||

QY 901 tccgctgctgatttgaacgctgctgctgcaagcgcgagcgtatatttggagcgttac 960
Db 297009 TCCGGTTCGTTATTTGAACCGCGGATTTACCAAGCGCGGACATTAATTTGGAGCTAC 296950
|||||

QY 961 cacaatcaagttccgcttcatcgagaagaagcgcgagcaagaagctgtgcgtgtgctg 1020
Db 296949 CACATTCAGATTCCTGTTATCGAAGAAAGCGCGCAAAAGCGCTGCGGTTGCG 296890
|||||

QY 1021 cgcgagccggaacaatactcatcaacgcgcacactctgcgcattctctaaatacaaa 1080
Db 296889 CCGCAGCGCGGACAAATATCTCATACGCGCTACAAACCTCGGCAATTCGTAACAA 296830
|||||

QY 1081 ctctcaagttcaacgaagcgcgtcaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1140
Db 296829 CTCTTCAGTTTCAACACAGCGCTCAACGCGCGGACCGCGCATGTGTCGATGTACT 296770
|||||

QY 1141 tatgagcgcgtaatgcccgttgagacatcctgctactctgtcttctgcgagtaattcgtc 1200
Db 296769 TACGAGCGCGTGTATGCCCTTGATATCTCCACACCTGCTTTGCGCGATTAATGTC 296710
|||||

QY 1201 ggcgataccgacagcgcgagcgttgggttgcttgatggaattgcaagaagaagcgtcgt 1260
Db 296709 GCGGATACCGACAGCGCGGACGAGATTGGTTGCTGGAATTTGACGAGAAAGCTCGCT 296650
|||||

QY 1261 ttgtagcgtctgctgcccgggcaataacgaatacagccgcgtgtgtgcgaagtgctg 1320
Db 296649 TTGTGACGCTTCGCTCCCGGCAATATGCAATACGCGCGCTGTGCGCAAAATGCTG 296590
|||||

QY 1321 gaaacccattgagaagaagcgtga 1344
Db 296589 GAAACCATTTGAGAAAGGCTGA 296566
|||||

RESULT 5
NMA372491/C 311321 bp DNA linear BCT 04-DEC-2000
LOCUS
DEFINITION Neisseria meningitidis serogroup A strain Z2491 complete genome;
ACCESSION AL162754 AL157959
VERSION AL162754.2 GI:7379424
KEYWORDS
SOURCE Neisseria meningitidis Z2491.
ORGANISM Neisseria meningitidis Z2491.
bacteria; Proteobacteria; Beta subdivision; Neisseriaceae;
Neisseria.

REFERENCE 1 (bases 1 to 311321)
AUTHORS Parkhill, J., Achtman, M., James, K.D., Bentley, S.D., Churcher, C., Klee, S.R., Morelli, G., Basham, D., Brown, D., Chillingworth, T., Davies, R.M., Davis, P., Devlin, K., Felwell, T., Hamlin, N., Holtroyd, S., Jags, K., Leather, S., Moule, S., Mungall, K., Quail, M.A., Rajandream, M.A., Rutherford, K.M., Stammers, M., Skellern, J., Whitehead, S., Spratt, B.G. and Barrell, B.G.
TITLE Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
JOURNAL Nature 404 (6777), 502-506 (2000)
MEDLINE 20222556
REFERENCE 2 (bases 1 to 311321)
AUTHORS Parkhill, J.
TITLE Direct Submission

		/protein_id="CAB83990.1"	
		/db_xref="GI:7379428"	
		/db_xref="SPTREMBL:Q9UVT7"	
		/translation="MPKLHMEFLYSGNAGRSNIEVHD IQEAVCDYD REAVALPKAAFC	
		DADKIHIDMQVWEADVADIIVASEPRTKMPSENPRIYFANYGRAQGLAEHAHA	
		GLPEAFPAEKOKALQTLITDSYYOQHMDNLKDYNLLADLRIGNHIRTLPDPHG	
		PATGFGGYLPI"	
		3782..3791	
misc-feature		/gene="NMA0703"	
		/note="Core DNA uptake sequence: gccgtctgaa"	
		/label=DUS	
		4083..4505	
gene		/gene="NMA0704"	
		4083..4502	
misc-feature		/gene="NMA0704"	
		/note="Pfam match to entry PF01242 PTPs, 6-pyruvoyl	
		tetrahydropterin synthase, score 10.80, E-value 1.7e-06"	
		4083..4505	
CDS		/gene="NMA0704"	
		/note="NMA0704, len: 140 aa; similar to many hypothetical	
		proteins e.g. SW:YB90_HAEIN (EMBL:U32798), H11190,	
		Haemophilus influenzae hypothetical protein (141 aa),	
		fasta scores: E(): 5.1e-32, 58.0% identity in 138 aa	
		overlap. Shows very weak similarity to eukaryotic	
		6-pyruvoyl-tetrahydropterin synthases e.g. SW:PTPS_RAT	
		(EMBL:M7850), pts, Rattus norvegicus 6-pyruvoyl	
		tetrahydropterin synthase precursor (EC 4.6.1.10) (144	
		aa), fasta scores: E(): 0.0062, 27.3% identity in 154 aa	
		overlap. Contains pfam match to entry PF01242 PTPs,	
		6-pyruvoyl tetrahydropterin synthase	

Query Match	87.7%	Score 1179..2;	DB 1;	Length 311321;
Best Local Similarity	92.3%;	Pred. No. 1.4e-286;		
Matches 1241;	Conservative 0;	Mismatches 103;	Indels 0;	Gaps 0;

OY	1	atgattaataatcaaaaaagtcctaatttgcacatcgcyggcgaaccgagcaaatcatt	60
Db	55476	ATGATTAAATCAAAAAAGCTTAACCTGCCCATCGCGGCAGACCGGACAATGTCATT	55417
OY	61	tatgacgccgcgcgatccatcacgaagtgcgcttgcttgcggaagaatatgfcgcatgcg	120
Db	55416	TATGACGGCGCCGTCATTACCAGACTCGCGTGCTTGCGGAAGAATAATGCGCGGTATGCGC	55357
OY	121	ccctcgatgaanaatcaaaggaaagtlgaagccgltcaaaaagaagccaadtgcgtttgaagac	180
Db	55356	CCCTCGATGAAGTCAGGAAGGAGGCGATGCGCCGTCAAAAGGCCAAGTCGTTGGAAGAC	55297
OY	181	aaaaagaatccgggcgttagtatctactgcgcgcggttcagagcaaatcgcgctattcac	240
Db	55296	AAAAAGAAATCCGGCGGTGCTTTAACCCGCGCGGTTCAGGCCAAAATGCGCCGCATCAT	55237
OY	241	cgtgcgcaaaaagcgcgtacctcagtcagtcgctgcatgcatgcgcttgaagaagcaagcagaatc	300
Db	55236	CGCGCGCAAAAAGCGCGTACTTTCAGTCGGTCCGATGTCGCCGTTGAAGGCAACGAGAAATC	55177
OY	301	gagtlcgaaactcgtactcgttaagcgcgtlbgcaaaatlbgacagcgaaaaagtlgcgcgc	360
Db	55176	GAGTTCGAACGCTACGGCCCCGAGACGCTTGGAACAACTTAAGCGCGCGAAGTAGCGCGCGC	55117
OY	361	aacctgattcaatcagcttatgactatgactgcgcttcgcaaccgcttcgcttaagcaaatccct	420
Db	55116	AACCTATCCAATCCGCTTGTGTGCACTCGCGCTCACCCGCTTACCCGCTTACCAAAATCCCT	55057
OY	421	gccgtatgatgcgagccgttcgcacatcttcgtcgaatlgatgagaaaccaatccgctgcct	480
Db	55056	GCCGTGATGCCCGAGCGCTTGCCATCTTTCGTAATGCCATGAGACCAATCCGCTGCGC	54997
OY	481	gccgacctaacggtlcalcatcaaaagaaagccgcaagactcaaaagcgcgcttgtgta	540
Db	54996	GCAAGACCCGTGGTGTGGATCAAAAGAACCCGCGAGGATTTCAAGACGAGGTTTGTCTGTGA	54937
OY	541	ttgagcgcgctgacgcaagcgttaaatccatgtgtgttaaagcagcagcgagcagcgtgcg	600

Db	5436	TTGAGCCGTTTGACCGACGCGTAAATTCATGTGTGTAAAGCAGCTGGCCGACAGCTGCCG	54877
Qy	601	tcctgaaaatgctgcgcacatactgaacacatgaatttggcggcccgcatctgcgcgcttg	660
Db	54876	TCTGAATATGCTGCCACACTTCGAACACATGATTTGGCGGCCGCACTCGCGCGTTTG	54817
Qy	661	atggcgcgcacatcatcttcatctcgagccagtcggcgcgataaaccgctgtgacacatc	720
Db	54816	AGTGACAGCAGCATTCATTCTATTGTAGCCGGTGGTGTCAAAACAACCGTTTGGACATTC	54757
Qy	721	aattataaagcgtatgtgtctatcggacgctttgttcgtaacaggccgtctgaataccgag	780
Db	54756	AATTAATCAAGATGTAAATTTGGCCATCGGACGCTTTTGTGCAACAGGCCGCTCAACACGAG	54697
Qy	781	cgcgtgtgtgccttggcggcgctgcacatgaataaacccgcgcctcttgcgacaccttg	840
Db	54696	CGCGTGTATTTGCTTGGGTGTCTCTCAAGTCAACAAACACGCGCTTGTGCGTATCCGTTTGG	54637
Qy	841	ggctgcgaagtgctctcaacttaccgcgcgcgcgaattggcttgcgcgcgaacccgcgtgat	900
Db	54636	GGTCCGAAGAATCGCAATATTACTGGGGGGGATTTGGTTGACGACGACACCGCGTGATTT	54577
Qy	901	tcgcgttcgctatgtgaacgctgctgatctcaaacaggcgcgcatactatcttggagcgtac	960
Db	54576	TCCGCTTCGGTATTGGAACGGCCGCGATTACACAAGCGCGCACGATTTATTGGGACGCTTAC	54517
Qy	961	cacataagatcttcgcttatctgaagaaggccgaagaagaagcgtgttcgctgtgttcg	1020
Db	54516	CACATATCAGATTTCCGTTATTCGAAGAGGCCGCGACAAAGCTGTTGCGCTGGGTTGCG	54457
Qy	1021	ccgcgaagccgaacaaatctcactcaccgcgcacacctctgcgcattctctctaaaaaacaa	1080
Db	54456	CCGAGCGCGGACAAATACTCCATCAACGGGTAGACCCCTCGGCCATTTCTGTAACAAACAA	54397
Qy	1081	ctcttcaagttcaacgaacgcgtcaacggcgcgcgcgcgcgcgtatgtatccgatctggcact	1140
Db	54396	CTCTTCAAGTTTCACGACAGCCGCTCAACGGTGGCGACCGCCCATGTGTGCCATTGGTGTACT	54337
Qy	1141	tatpaggcgctaatgcccgtttgacatccgcctaccttggcttttggcgatattatcgtc	1200
Db	54336	TACAGCGCGTAATGCCCCCTAACATCTCGCCCTGCTTTGGCCGATTTAATCTGC	54277
Qy	1201	ggcgataccgaacagcgcgcgcgcgttggcttggcttgcgttgaatttggacgaagaagacctgcct	1260
Db	54276	GGCGATATCCGACAGAGGGCCCAAGCATTTGGTTGCTTGGAAATTTGACGAAAGACCTGCCT	54217
Qy	1261	tttgtgaagcttcgtctcgcgcgggacaatacgaatacgcgcgcgctgttgcgcaagctgctg	1320
Db	54216	TTGTGCGAGCTTCGCTCTGCCCCGGGCAAAATCGAATAACGGCCCGCTGTGTGCGTAAGGCTGTG	54157
Qy	1321	gaaaccatitgaagaagaagcgtga	1344
Db	54156	GAATCCATTGAGGAAGAGAGGCTGA	54133
RESULT	6		
LOCUS	A96094	1344 bp	DNA
DEFINITION	Sequence 127 from Patent WO9924578.		
ACCESSION	A96094		
VERSION	A96094.1	GI:6779939	
KEYWORDS			
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1344)		
AUTHORS	Pizza,M., Scariato,V., Rappuoli,R., Grandi,G. and Masignani,V.		
TITLE	Neisserial antigens		
JOURNAL	Patent: WO 9924578-A 127 20-MAY-1999.		
	PIZZA MARIAGRAZIA (IT); SCARLATO VINCENTO (IT); RAPPUOLI RINO (IT);		
	CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VERGA (IT)		
FEATURES	location/Qualifiers		
source	1..1344		

/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 323 a 347 c 368 g 289 t 17 others
ORIGIN

Query Match 86.7%; Score 1165.4; DB 6; Length 1344;
Best Local Similarity 91.2%; Pred. No. 2.9e-283;
Matches 1226; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

Qy 1 atgataataataaanaaggtcttaaatctgcccacatcgcgggcagcagcagcgaagtcatt 60
Db 1 ATGATTAAATCAAAAAAGCTTAAACCTGCCATCGCGGAGACCGGAGCAAGTCATT 60
Qy 61 tatgacggcccgccattccgaagtcgcttgcttgcgcaagaatatgctgcgcatcg 120
Db 61 TATGACGGGCCCGGTCAATACCGAAGTCGGCTTGCGCAAAATATGCGGATATCGC 120
Qy 121 ccctcgtatgaatacaagaagtgaaagcgtlcaaaaaaagcccaagtcgtcttgaagac 180
Db 121 CCTCGATGAAGTCAAGAGGCGATGCCGTCACAAAAAAGCCAAAGTGTGTTGAGAC 180
Qy 181 aaaaagaatcccggtgagttattactgcccgcgcttcagggcaaatcgccgtatcac 240
Db 181 AAAAAAGNATCCGGCGGTGGTGTACCGCCCGCTTACAGGCAAAATCGCCGCATCAT 240
Qy 241 cgtgacgaanaagcgcgtactcaagtcagtcgcttgcttgcaagcgaacgaacac 300
Db 241 CGCGGCAAAACGGGCTACTTCACTGCGTGTGATTCGCTTAAGGCAACGACGAATC 300
Qy 301 gagtcgaacgctacgtacgtacgtctgcaaaaattgagcaagcaaaaagtgcgcgc 360
Db 301 GAGTTCGAACCGCTACGCGCCCAAGCGTTGGCAAACTTAAGCGCGANGAANTNGNNG 360
Qy 361 aaccgtatcaatcagagctatagactgctgcacaccgctcgttcaacaataatcc 420
Db 361 AATCTGATCCATCCGCTGTGTGAGCTGCGCTGTANCCGCTGTCACGAATAATCCCT 420
Qy 421 gccgtagatgcgagccgcttcgcatcttcgtcaatgctgaatgcaacacatccgct 480
Db 421 GCGGTGATGCGGAGCGCTTCGCCATCTTGCTAATGCGATGGAACACCATCCGCTGCG 480
Qy 481 gccgacccctacggtcatcatcaagaagccgcggaagattcaacgcggtctgtgta 540
Db 481 GCAGACCCCTGTGTGATCAAAAGAGCCGCGANGATTTCAGACGANGTNGCTGTGA 540
Qy 541 ttgagccgctcagcgaagaatcaatccatgctgttaagcaagcagcagcagctgcg 600
Db 541 TTGAGCCGTTTGAACGAGCGTAAATCCATGTGTGAAGCGCTGCGAGAGCTGCG 600
Qy 601 tctgaanaatgctgcacatacgaaacacatgaattgagcgccgcgcatcctcgcgct 660
Db 601 TCTGAANAATGCTGCCAATCGAAACACATGATTCGGCGCCCGCTCCGGCGGCTTGG 660
Qy 661 agtggacgcacattcatctcagcagctgcgcgagcaataaaaacgctgtgcacatc 720
Db 661 AGTGGACGACCATTCATTCATTGAGCGGCTGCTGCAAAACAAACCGTTTGGACGATC 720
Qy 721 aatttaagaagctgattgctatcgagcttggctcgtaacagcgctcgtgaatacag 780
Db 721 AATTATCAAGATGATTAATCCATCGGACGTTTGTTCGAACAGGCGCTGTGAACAGAG 780
Qy 781 cgcgtgctgcttgagcgccctgcagatcaacaacacgcgctctgcgtacccgtttg 840
Db 781 CGCGGATTCGTTGGTGTGCTCTCAAGTCAACAAACACGCGCTTGGCTACCGTTTGG 840
Qy 841 ggtgcgaaggtctcctaactaccgcgcggaatggttgaacgcggaacaaacgcgtgatt 900
Db 841 GGTGCGAAGGTCTCCGAATTTACTCGGCGGAGATGTTGACGACGACAAACCGCGTATT 900
Qy 901 tccggttcgcttgaacggtgcgattgcacaagcgcgcatgatatcttggaacgctac 960
Db 901 TCCGGTTCGCTTGAACGCGGCGGATTTACCAAGGCGCGCATTAATTTTGGGACGCTAC 960

Qy 961 cacaatcagattccgtatcgaagaagcgcgcaagaagcgtgctgcgtggttcg 1020
Db 961 CACATTCAGATTCCGTTATCGAAGAAGGCCCGCAGCAAGAGCTGTTCGGCTGTGCG 1020
Qy 1021 ccgcagccggaataactcatalcaacgcgcacacactcgcgcatttccttaaaaaa 1080
Db 1021 CCGGACCGGCAAAATACATCCATCCACGCGTACGACCTCGGCATTTCCTGAACAA 1080
Qy 1081 ctctcagttacagacgcgttcacagcgcgcgccgcgcgcatgtatcagtcgcgact 1140
Db 1081 CTTTCAAGTTTCAGACAGCCCTTCAACGGTGGCGACCGCCCATGATGGTACT 1140
Qy 1141 tatgacgctaatgcgcttggacatccgctacacttccttcttgcgcatlaacgctc 1200
Db 1141 TACGAGCGCGTATCCGCTTAACATCCCTGCTACCTCTTTGGCGCATTTAATCGTC 1200
Qy 1201 ggcgaatacgcagcgcgcagcgttggctgcttggaattgacgaagaagacccgc 1260
Db 1201 GCGCATACCGACAGCGCGCAACATTCGCTTGAATTGGACGAAGACCTCGCT 1260
Qy 1261 ttgtgacgttcgcttcgcccgggcaaatagacagccgctgttgcgaagtgctg 1320
Db 1261 TTGTGACGCTTCGCTCCCGGCAATACGAATANGCCCGCTGTGCTAAGTGTCTG 1320
Qy 1321 gaaacatlgagaagaagagctga 1344
Db 1321 GAACCTTTGAGAAGGAAGCTGA 1344

RESULT 7
AE006171 10563 bp DNA linear BCT 08-MAR-2001
LOCUS
DEFINITION Pasteurella multocida PM70 section 138 of 204 of the complete genome.
ACCESSION AE006171 AE004439
VERSION AE006171.1 GI:12721691
KEYWORDS
SOURCE
ORGANISM Pasteurella multocida.
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

REFERENCE
AUTHORS 1 (bases 1 to 10563)
May,B.J., Zhang,Q., Li,L.L., Paustian,M.L., Whittam,T.S. and Kapur,V.
TITLE Complete genomic sequence of Pasteurella multocida, PM70
JOURN. Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)
MEDLINE 2 (bases 1 to 10563)
21145866
REFERENCE
AUTHORS Zhang,Q. and Kapur,V.
TITLE Direct Submission
JOURN. Submitted (24-OCT-2000) Department of Veterinary Pathobiology,
University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN
55108, USA

FEATURES
source
1..10563
/organism="Pasteurella multocida"
/strain="PM70"
/db_xref="taxon:747"
94..1605
/gene="PM1326"
94..1605
/gene="PM1326"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK03410.1"
/db_xref="GI:12721692"
translation="MWIDVWRODCLILKLNKSGSFDPVPLKNVDLYIAGEVHCLV
GENGAKSTLCKIAGIYPPROGEMFYGKSYEPKTVKRAOSLIGTHOELMIVPOL
SVLENIFENKRYFCGKMDQTMHKKQOIISELELDIDPPAKRADLSIAQQQWEI
AKAVSERKVIITFDEPTSSISKNTQTLFKIIHQJKEKNVAVIYSHRDEROYIADQ
VTVALRDGIRGTMAVYRDTSPDEIVKLMWGREINFAEYHHRHVIPTREVILVESLJTNKI

gene
CDS
1598. : 2641
/gene="rbsc_3"
/note="PM1327"
1598. : 2641
/gene="rbsc_3"
/codon_start=1
/transl_table=1
/product="RbSc"
/protein_id="AAK03411.1"
/db_xref="GI:12721693"
/translation="MNRVCMTRKIRFESKIGIITLLIMHIGMSLSDFVLTNRNIIN
ILLOVSTICVSVGMTIYIILTGIDLSVGSITVALSKVGLFTHWGMDWLGEMSSFS
VLMIVISAIIATIFVIGILGYYVNGVLIVYKVTPEITLGMGIAGLALTIIDGKTI
VNPEDTRFNGRRIKALTEFETPALPVPIAIIILVVVSFVYLQTMGRQRIYALGNRE
AYRSGINIEKLEIKAYETINAGLAIIGALINAQPIAGNGELDIATVIGES
TSLMGGSVSVSTIGALIMGVILONGTLLNANTSYLQRIIGIIVITLAVPLDQLRREE
VSTRKRKRIFPRE"
3187. : 4527
/gene="PM1328"
3187. : 4527
/gene="PM1328"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAK03412.1"
/db_xref="GI:12721694"
/translation="MITIKGILNPIIGSSPEQVIRDNAITEVALLGEEYVGNRPSNK
VRGCDVKKQVLEFEDKNKPGVVEFTPAPSGVTALIRGAKRVLOSIVIKIEGEOITFE
EKTEEDNQLDTEQVRONLQASGLMTALRTPEFSVPADATPVSIFVAKMPTNPLC
ADPOVIVYOOSAOAEAGLETVLSRLEHGAVYLCKANASIPSPINAMDVEFNGHPHA
GLSGTHHFDIVSATKFEVWYINQDVIYAVGKLTPTGELDSRVSLAGQVKNRPVY
RYVIGANLSOLTANEMVDGENRVIISGVSLSGAKAAGPVYILGRYALQVSYLEGRERE
FLKMDVGNANKYSLSTRYVLGHFSKLEFNFTALNGGERAMVPIGATERVMPDITPTL
LFLMDIANGDTSQAALGCLTEIDEDLALCTFVCPKNEYGAPILRQALDKIEKG"
4530. : 5762
/gene="nqrb"
/note="PM1329"
4530. : 5762
/gene="nqrb"
/codon_start=1
/transl_table=1
/product="Nqrb"
/protein_id="AAK03413.1"
/db_xref="GI:12721695"
/translation="MGLKHFEKTEPPAFIPGGYKRYVLAIFEATATFIVTPGYTHKA
SHVRDADSKRMVIVNLSIFPAMFYGMVNGAQAIIATDAGLTDOOAIISGHMOYALS
HALGADITLSGMAKSKMLIGARTYPLTYLTITLVGGFWELRVAMVRKHEINCEPEVTS
ILALIVPLPLIQOALATTFGVVAAKEIFGGVAKNFEMNPALAAAELEFAVPAQIJS
GDVWVAADSESGATLQSOMAVGGEAGLKHVVTGQPIITMDAFITGNIIGSIGVSTPA
LIGCAIVAVIARIASWRIIAGVMIGMITSALFNIGSSTNPLFAMPWYWHFVIGGGA
LGFEPMATDPVSAFTNKGKMWYGILIGAMAVLIRVYNPAYPEGMALILFANLFAVY
FDYLYVQANIKRRRARAY"
5755. : 6537
/gene="PM130"
5755. : 6537
/gene="PM130"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAK03414.1"
/db_xref="GI:12721696"
/translation="MFRNKDSVGGTLVYIILSLACSIIIVAGSAVLKPTOIEQKEID
KONNIIISVAGILOPTFRNSEIKTIYANNIEARLVDLNTGDFAPAGCGPAAKAKVNP
ESVALSEDDVAGIRVRANFAEVYIYKNDAGQTOVTVDFPYGGLSMYGMFVSQVD
GMVNGITTYDQGETPGIGGEILINPKMOAOPGKRLTYTDNOYGLVYKGAASNAEHG
IDAISGSLTISNGVNNSEFRWLGQKGFPLAKIAGVYIANG"
6530. : 7156
/gene="PM131"
6530. : 7156

gene	/gene="PM1331" /codon_start=1 /transl_table=11 /product-"unknown" /protein_id="AAK03415.1" /db_xref="GI:12721697" /translation="MADTFKKLGLILSPVMDNNPINALQILGICISALAVTFLKEPAIYM TPAIVFPTAFSLNLSLTIRNYIPNSRITTYOMAILASIVLYNDQILLRAVYAGSKOLS VFVGLITNCITVMGRAPAMKSQPLESFYVDIGNLGTGAILLYSAFIELLGSGKL EGMTVPQTIDDGWYOTNGLLPLLAFAPIIGHIWIIGIKLPNOVER"
CDS	7160..7756 /gene="ngre" /note="PM1332" 7160..7756 /gene="ngre" /codon_start=1 /transl_table=11 /product-"Ngre" /protein_id="AAK03416.1" /db_xref="GI:12721698" /translation="MEHYISLPEVKSVEFIENMALFELQCETPLAVSKYSTARGICIA VIVVIGIAPVNOLVYSFILKDSALVOGIDLSEFLNTPTTGVAALVOLLBEMLDKFE PALYNALGLEPLITVNCALIFGVGSFMVRDPTFVSYYVGIGATGMMAIYVALGIT TEKMRYADVPAQLRGITGFITVIGLMALGFMSPSGIOL"
gene	7772..8995 /gene="ngr6" /note="PM1333" 7772..8995 /gene="ngr6" /codon_start=1 /transl_table=11 /product-"Ngr6" /protein_id="AAK03417.1" /db_xref="GI:12721699" /translation="MEITRTGIMFTYIVIALAYIIIFAKSKLVNSGDITFEINDPSK AIHPRAGCKLICALASOGIFVSSACGGSCCGCITYKTVEGGDILIPLESHISKREPA KEGYVLSGCVANPKSMKYALEPVEFCVKRWECTVISNDKATFIKELKLAIPGEERP FRAGYIQIEANPTVYAKDFDIPEEYHEHDWKYNLMRVYSKVDENIIRAYSNASYPE EKGIIMLVNRITPPPNPDAPQGQSSYIWSLKAGDKVTIISGPFEEFFKEDNAEVY EVGGAGMAPMRSHIEDQLRLKRSKMSFWGARSKRMETFYDEDDTLQAENDFPW HVALLSDPQDNMGDTGTFIHNVLENYLNKHDEAPDCEYMCGPIMNASVIKMLKD LGVEDENILIDDFGC"
CDS	9164..10204 /gene="apbe" /note="PM1334" 9164..10204 /gene="apbe" /codon_start=1 /transl_table=11 /product-"Apbe" /protein_id="AAK03418.1" /db_xref="GI:12721700" /translation="MKHIINSLITIGFCSLCILACHKSPREVWTITGTMGTTSVSKYID DGOATIKPLEMHGOIEELIKEVNNKNKSTYLSTSELSONFOQIDHPVESADLAFLV KEATLHGVTGEGALDYTVGPVYINIMGFGEPEROPNETLEOYEKRBRAMWGIEKMLLE OCGRKYIAKTVPOLXYIDLSIAKGCVDPVAHAHLAKGTANYLYVELGETEINAKGNIE GDMOWIAIEKPENFDGSRAYSQITIGLKDFSMAITSGNTRNFTEENGRRPSHEDTKTCIP IEHRLASITVLADSTMTADLGISTGLVIGEERALEVAERYNLLVYLISKTEKFTEAM SSAEFKLNEQH"
gene	10214..10468 /gene="PM1335" 10214..10468 /gene="PM1335" /codon_start=1 /transl_table=11 /product-"unknown" /protein_id="AAK03419.1" /db_xref="GI:12721701" /translation="MOPLITRGIFILITAGMAICYITIOKRTIAGSCGGISALGLAKKY CDCEPCDNLKAKLEAGDEEAODEYNAKFAKNBPHTYEVK"
CDS	
BASE COUNT	3052 a 1783 c 2451 g 3277 t
ORIGIN	

REMARK	The H. influenzae sequence has been updated by R. Fleischmann. New database matches have been assigned, product names have been improved, and a number of frame shifts have been corrected. We gratefully acknowledge the work of Tatusov et. al. We have incorporated their annotation into the /notes fields of the corresponding H. influenzae genes
REFERENCE	5 (bases 1 to 13299)
AUTHORS	White, O., Clayton, R.A., Kerlavage, A.R., Fleischmann, R.D., Peterson, J., Hickey, E., Dodson, R. and Gwinn, M.
TITLE	Direct Submission
JOURNAL	Submitted (28-May-1998) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
REMARK	The whole genome was shifted by 588 nucleotides for a new start
COMMENT	On Sep 30, 1996 this sequence version replaced gi:1220812.
FEATURES	Location/Qualifiers
source	1..13299
gene	/organism="Haemophilus influenzae Rd"
CDS	/db_xref="taxon:71421" complement(85..1455) /gene="HI0161" complement(85..1455) /note="similar to SP:P06715 GB:M13141 PID:146248 PID:46637 GB:U00096 percent identity: 74.17; identified by sequence similarity: putative"
gene	/codon_start=1 /transl_table=11 /product="glutathione reductase (gor)" /protein_id="AAC21833.1" /db_xref="GI:1573119"
CDS	/translation="MTKHYDYIAIGGSGGIASLNRAASYGKCAIIEAKHLGTCVNVCCVPKWMFEGAHIEANNVADPGGEVDEKPFDSKLESROAVISRHTSYNVLAKNNITVNGSGKFPVDAITETVLTADGKEQVPAADHILATGCGRPAPRNPKEGEGTDSDFGPAITELPKRAAVTACGTATVETSGVLSICVETHLVRRHAPRNPDLIVETLVEVAQDGIQLHTNSTPSEIYNADGSLTVRCDSGDVTVDCVMAAGRPPTDKTG LNEAGVETNEHGVKVDKQYQNTNKGIIYAVGDIENGIELTPVAAAGRRLSERLFNN KPEYLEDYSLVPTVYSHPIGVGLTEPOAIEQYGAEVNVKVSFTAMTAVTQHR OCKMKLVCVGDEKVVGLHIGFVDEMIGFAVAIKMGATKADFNTVAIHPTGSE EHYTMK"
gene	complement(1671..2270) /gene="HI0162"
CDS	complement(1671..2270) /gene="HI0162" /note="similar to GB:U00096 PID:1773118 PID:1786637 percent identity: 25.82; identified by sequence similarity: putative"
gene	/codon_start=1 /transl_table=11 /product="11poprotein, putative"
CDS	/protein_id="AAC21834.1" /db_xref="GI:1573120"
gene	/translation="MALVYKIKTLSSVGLIANTFLACGAAQSNIIAFTPPASASNVNRTAAVSVTTKDSRAIOETASTTKHEGLTKLNAPSVYTDLFQVMOONLISKGRVQ LNSNAWVTVDFRFGTOVEGNLRKRLTKTQATVYVOCAGKSYNKSFNVTSHQSG VFNAGNDEIHKVLSQTFNDIYNNIYDOEVAAMAINQYSN"
gene	2364..2675 /gene="HI0163"
CDS	2364..2675 /gene="HI0163" /note="similar to PID:663269 percent identity: 46.53; identified by sequence similarity: putative"
gene	/codon_start=1 /transl_table=11 /product="transcriptional regulator, putative (bolA)" /protein_id="AAC21835.1" /db_xref="GI:1573121"
CDS	/translation="MSTQQLIEOKIQEOPHFLAENESHLHNSNGSESHFKCVIY SADFKNIRKQVQRHRIYQLINELNHSIALALHLFPEEKAKQNETVPSTKACAGIG R"
gene	2966..4309 /gene="HI0164"
CDS	2966..4309 /gene="HI0164" /gene="HI0164"
gene	/note="similar to PID:663269 PID:663270 percent identity: 64.96; identified by sequence similarity: putative"
CDS	/codon_start=1 /transl_table=11 /product="NADH:ubiquinone oxidoreductase subunit A, Na translocating"
gene	/protein_id="AAC21836.1"
CDS	/db_xref="GI:1573122" /translation="MITTKGIDLPIAGKPAQVYIHSGNAVNOVAIIIGEYVGRPSMK VREGDVVKGQVLFEDKKNPQVIFTPAASGTTPAIRGKERVLOSIVINVEGDEKITF AKSTVEQLNTLSEEVKONLIESGLMTALRTPESVPSIESPASISFVNMOTNPILA ADPSVLEYSODFTNGLVLSLRFPSKPLHCKADSNIPTADLENLQIHDTGTGPH AGLVSTHAEFDTPVGIQKTVMHINQDVAVKGLFTTGELYSERVALQSVIAEGNEK VFTTIGANLSQLTQNELSAGKNRIVISGVLCQIADSHDYLRVALQSVIAEGNEK EEEGWIMPOANKSVTRVYVGHGSKLFFNTTSENSGERAMVIGSERYEMPDILPT LILRLIDYGDPTGQAOELGCLDEDEDLALCSFVCPKYEGLSRQVLDKIEKG"
gene	4312..5547 /gene="HI0166"
CDS	4312..5547 /gene="HI0166" /note="similar to PID:663269 PID:663271 percent identity: 71.33; identified by sequence similarity: putative"
gene	/codon_start=1 /transl_table=11 /product="NADH:ubiquinone oxidoreductase, subunit B (nqrB)" /protein_id="AAC21837.1"
CDS	/db_xref="GI:1573123" /translation="MGLNLFEMKEPAFLPGKYSKLYPIFESITLLTYTPGVTHKN THVRALDSKRMKTIYFIALFPAIFGYMNVGNQALIPALMQLNLQILLANDHVALA SLSGLDLPANATWGSKMALGALIFGLIYLVFVYVCIWELLSFVNGHEVNEGMPVST ILFALIVPTLTLPMQALSGITFGLIYAKELFGVGNFNNMPLAAGALFEFAVPAQIS GDTVMADGFEFGATALSSOMSGOGALQHVTVGAPITMDADEVNLP.GSMGEVSTLA ILIGCAVYFETRIAMRLIAGVIGMIATSTLENLSEFNPMFSDWHPHVLGGFA LCMVKNATDPVSASTNTKRMWVGALIGVMAVILRVNPRYPSGMALILFANLFARI EDYIVQANIKRRARTNC"
gene	5540..6274 /gene="HI0167"
CDS	5540..6274 /gene="HI0167" /note="similar to PID:663269 PID:663272 PID:893413 percent identity: 54.36; identified by sequence similarity: putative"
gene	/codon_start=1 /transl_table=11 /product="NADH:ubiquinone oxidoreductase, Na translocating"
CDS	/protein_id="AAC21838.1"
gene	/db_xref="GI:1573124"
CDS	/translation="MAKFNKDSVGGTILVVLISVCSIIYAGSAVWLKPADEOKILL DKONILNTVAGLQENTNFKETAFKLEPFEVDLAVGEYQADDGQAIPADADKAR ITRSKTTEVYLVKDEGOTQOVLILPYGGLMSWAGLVSOPDGNTINGITYOHG ETPGAGGELIENPNWASLFFGKKRLFDSOHOPAIRIVYGOAQDEHSIDGLSGATLTNG VGGTEVWFPSKDSGFYLERKLSGAN"
gene	6274..6900 /gene="HI0168"
CDS	6274..6900 /gene="HI0168" /note="similar to PID:663269 PID:663273 PID:893414 percent identity: 74.88; identified by sequence similarity: putative"
gene	/codon_start=1 /transl_table=11 /product="NADH:ubiquinone oxidoreductase, Na translocating"
CDS	/protein_id="AAC21839.1"
gene	/db_xref="GI:1573125"
CDS	/translation="MSGKTSYKULLAPIAKNNPDIALQILGICSAVATTKLEFAPYV AIAVTLVTGSLNLFILIRNYIPNSIRIYQALILASIVLVQDILAAVYGLSKOLS VEVGLITNCLVGRAEAFAMKSPVESFGLGNGIGLVSMILIVAFPELLIGSGKL FGMTIETIONGCVQANGFLILAPSAFIIIGVYIWMGLRTWKPEQCK"
gene	6904..7500 /gene="HI0170"

CDS

6904..7500
/gene="H10170"
/note="similar to PID:893416 PID:893415 percent identity: 84.34; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH:ubiquinone oxidoreductase, Na
translocating"
/protein_id="AAC21840.1"
/db_xref="GI:1573126"
/translation="MEHYSLSFKAVPIEMNALSFFGMCFTFLAVSKVSPAFGLGIA
VTFVGIAPVMDLIITANVLKENALLIEGVDSFLNFTFTGIVLQVILDMVDKFM
PSLVNAGLIEPLIAVNCALIEGVSEFMVDYNEPESIIYVGFSGSLGMALIVLAGL
TEKMYADIPAGIKGLGITFISVGLMALGFMSFGIOL"

Query Match 40.9%; Score 550.2; DB 1; Length 13299;
Best Local Similarity 64.1%; Pred. No. 6e-128; Mismatches 478; Indels 6; Gaps 2;
Matches 863; Conservative 0; Mismatches 478; Indels 6; Gaps 2;

QY 1 atgattaaatcaaaaaggtcctaatactgcacatcgcgagcgagcgagcaagtcatt 60
DB 2966 ATGATTACAATTAGAAAGGTTTGATCTCCCAATTGGGGGAAAACGACACAACTATTC 3025
QY 61 tatgagcgccgccccttaccgaagtcgcgttcgttggcgagaatatgctgcgacgcgc 120
DB 3026 CATAGCGGCAACGCTGTGATCAAGTTGCGATTCTAGGTGAGGAGTATGTGGGATGCGT 3085
QY 121 cccctgataaaaatcaagaagtgtaagccgtcaaaaagcgcaagtgctgttgaagac 180
DB 3086 CCTTCATAGAGGTGGCGGACGAGTGTTGTGAAAAAGGTCAAGTACTTTTGAAGAC 3145
QY 181 aaaaagaatcggggcgtagtattactgcgcgcgttcgaagcaaatccgcctattcac 240
DB 3146 AAAAAAAATCTGTGTATTTTACAGCCCTGCAAGCGGTACCATCATGCATTCAT 3205
QY 241 cgtgagcaaaacgogtacttcagtcagtcgtgatctgcgttgaagcaagcaaatc 300
DB 3206 CGTGCAGAAAACGCTGTATTACATCTGTGTTATATGATAGAAAGGATGAAAAATC 3265
QY 301 gattcgaaagcttacttcgaagcgctgcaaaatctgagcaagcaaaagctgcgcgc 360
DB 3266 ACTTTGGCAAAATATATGACAGACAACTTGAATACACTTTCTTGACAAAGTAAACAA 3325
QY 361 aactcgattcaatcaggtctatgactgcgttcgcacccgcttcgaagcaaatccct 420
DB 3326 AACTTGATGATCTGGCTTATGAGCGGCAATGACTGCTGCATTAGTAAAGTTGCA 3385
QY 421 gccgtagatgcgcgagccgttcgcacatcttcgtaagcgatgagcaacatccgctgct 480
DB 3386 TCCATTGAAAGTGAAGCATCTCTATTATTGTAATGCGATGATACCAATCCATTACCA 3445
QY 481 gccgacctacgctacatcaaaagaagcgcgcaagacttcaaaagcgctgttgta 540
DB 3446 GCCGATCCCTCTGCTGCTATTAAGAAATATGCCAAGACTTCACTAACGGTTTAAACCTTA 3505
QY 541 ttgagcgccctgaacgaagctaa---aatccatgctgtaaaagcagcgcgcgcgcgc 597
DB 3506 TTAAGTCTCTCTATTCCTTCAAAACCATTAACCTTATGTAAGACGCGATTTAATATT 3565
QY 598 ccgctctgaataatgcgcacatatcgaaacatgaatttggcgccgcgactctgcgcgc 657
DB 3566 CCAACTGCTGATCTTGAATAATTACAGATTCAGATTACTGCTGTTCACTCCAGCTGGT 3625
QY 658 ttgagtgagcacacatcatctcaatccagcgagcgcgagcaataaaacgctgtgac 717
DB 3626 CTTGTAGGAGGACCATATTCATTATTGATCCGTGCGATTCAAAAACCTGATGCGAT 3685
QY 718 atcaattatcaagcagctgattgctacgcagcttctgctlaacagcgctctgaatacc 777
DB 3686 ATCAATATTCAAAGATGTATGCTGTAGTAAATATTACAAACAGCGCAACTTATTC 3745
QY 778 gagcgctgctgtccttggcgctgcgaagcaaacacgcgcctctgctgacgct 837

DB 3746 GAACGAGTAATTTCTCTTGACAGTCCCTCAAGTGAAGAAACCTGTTAGTACGTACACA 3805
QY 838 ttgggtcgaaagtggtctccaacttaccgcgcggaattgttgagcggaacaacgcgtg 897
DB 3806 ATCGTGCTAATCTTTCATTAATTAACCCAAAATGAATTAAGTGCAGGTAAAAATCGGTA 3865
QY 898 attccgctgctgatttgaagcgctgcgatttgcacaagcgcgcatattatttggagcg 957
DB 3866 ATTCTTGCTAGTTCCTTTTGGTGTAATAATAGCAAAAGATTCAATGATTTATAGTGGC 3925
QY 958 taccacaatcagatttccgctatcgaagaagcgcgcaagaagctgttcgctggtt 1017
DB 3926 TATGCTTGCAAGTATGATGATCCCTGAAGTAATGAAGAAATTTTGGTTGGAT 3985
QY 1018 ggcgcgagccggaacaatactccatccgcgcgcacacactctggccatttcttaaaac 1077
DB 3986 ATGCGCAAGCCCAATTAATATTCGTGTGACTCTACCGTATTAAGCCACATTC--AGTAA 4042
QY 1078 aaactctcaagctcacacagccgtcaacgcgcgagcgccatggttaccgacgcgc 1137
DB 4043 AAATGTTTAATTTTACTACTTCAGAAAAATGTGGCGAAGCTGCATATGTCATTCG 4102
QY 1138 actatgagcgcgtaatgctgacacatccgcttaccctgtcttgcgcgattatc 1197
DB 4103 AGCTATGAGCGCGCTAATGCCATTTGACATTTTACCTATTAATTAATGATTAATC 4162
QY 1198 gtcgagcatccgcgacgagcgcgagcttgggttcttggaaattgagcaagaagcc 1257
DB 4163 GTGGCGCATACCGATGCGTGCACAAAGTATGATGTTCTGGAATTAAGCAAGAACTTG 4222
QY 1258 gcttltgcagctcgtctgcgcgcggaacaatacgaatacgcgcgcgttgcgcaagtg 1317
DB 4223 GCGTTATGTTCTTCTGTTTTCAGGCAAAATATGAAATACGTTTGGCTCAAGTC 4282
QY 1318 ctggaacacattggaaggaagctga 1344
DB 4283 TTAGATTAAGATTGAAGGAAGGTTAA 4309

RESULT 10
AF117331
LOCUS 5916 bp DNA linear BCT 13-MAY-1999
DEFINITION
Vibrio cholerae N16961 Nat--translocating NADH-ubiquinone
oxidoreductase enzyme complex, complete sequence.
ACCESSION
AF117331
VERSION
AF117331.1 GI:4808841
KEYWORDS
SOURCE
ORGANISM
Vibrio cholerae.
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE
1 (bases 1 to 5916)
Hase,C.C. and Mekalanos,J.J.
Effects of changes in membrane sodium flux on virulence gene
expression in Vibrio cholerae
Proc. Natl. Acad. Sci. U.S.A. 96 (6), 3183-3187 (1999)
JOURNAL
99179036
MEDLINE
2 (bases 1 to 5916)
Hase,C.C. and Mekalanos,J.J.
Direct Submission
TITLE
Submitted (31-DEC-1998) Microbiology and Molecular Genetics,
Harvard Medical School, 200 Longwood Ave., Boston, MA 02115, USA
JOURNAL
Location/Qualifiers
FEATURES
source
1..5916
/organism="Vibrio cholerae"
/strain="N16961"
/db_xref="taxon:666"
/note="biotype: El Tor"
34..1374
/gene="ngra"
34..1374
/gene="ngra"
/note="alpha subunit of NOR enzyme complex"
/codon_start=1

```
/transl_table=11
/product="NgrA"
/protein_id="AAD29962.1"
/db_xref="GI:4808842"
/translation="WITIKGLDLPDIAGTPSOVIVSDGKAIKKVALIGEEVYGRPTMH
VRGDEVKRAQILFEDKKNPGVKTSPVSGKVEINRGARKVLQSVLIEVAGDOQVTF
DKFERNOIASINRDAIKTOLVESGLMTAFRTPEFSKPAIDSTSEAFVYAMTDNPJA
AEPTVYINEOSEAFAGLDVLSALTGKVCYCKGSLPSOOPNVEEHYFDGPNHAG
LACTHMHFLVPSADHVAASINVDVIANIGOLFLNGELTORVYSLAGPVNKPRLVR
TVKASLEQULVDSEIIMPGEVRITSSVSGTKATGPHATLGRHLOVSVLREKDEKL
FGMAPGNKFSVTRSFUGHLEKGOVYNNMTTNTGSDRSMPVPIGNETKVMPLDMEPFL
LLRDLACDSDSVARLGALEDEDLALCTFVCPKCYEGQLRECLDIKIEKG"
1380. 2627
/gene="nqrb"
1380. 2627
/gene="nqrb"
/note="a subunit of NOR enzyme complex"
/codon_start=1
/transl_table=11
/product="NgrB"
/protein_id="AAD29963.1"
/db_xref="GI:4808843"
/translation="MGLKKEFDIEHHEPFGKHEKWFALYEAAATLEPTPGLVTKRS
SHVRSDVLKRIMVWLAVPFAMFWMYNAGQAAVLAUNHLYSGQLAAIVAGNMHY
WLTMLGSTMDSAGMSKMLGATYFLTYAVFVYGFWEVLFCMKVRHENVGEFV
VTSILFALIVPTPLPMOALGITRGVAVAKVEFGGTGRNFLPALAGRAFLFEAPYA
OISGLVMTAADYSGATATLSQMOAGAGALINNAIGOTITMDAFLGNTPGSIGEVS
TLALMGAAFLYIMKSIASRRITIGVMAIGLILSTLEPNVIGSDTNAMFNMMHHVLYG
GFARGFMTAIDPVASAFSTNSGRMAIGLIGVAVLIRVYNPATPRGMALITFANLE
APLFDHVVVERIKRRLARVCKO"
2614. 3387
/gene="nqrc"
2614. 3387
/gene="nqrc"
/note="gamma subunit of NOR enzyme complex"
/codon_start=1
/transl_table=11
/product="NgrC"
/protein_id="AAD29964.1"
/db_xref="GI:4808844"
/translation="NASNNDISIKTLFVYIALSLVCSIIYSAVAACLRDOKRENAID
KOSKILQVAGIEAKSKQIVLEFNKSTIEPLVDENGTDEVEDAANTDOKRAKEASE
SIKILAEODKAKIORRANGVVYLVKDGKTSKVLIPYHONGIAMSMTAVAAETDGN
TVSGLTVEQGETPGIGEVENPAMRAQVWGKLFEPNHPAIKIVYGAGPOGSEHGV
DGLSGATLTSNCGVQMTDFWLDMGFGPFLTKVRDGLN"
3387. 4019
/gene="nqrd"
3387. 4019
/gene="nqrd"
/note="b subunit of NOR enzyme complex"
/codon_start=1
/transl_table=11
/product="NgrD"
/protein_id="AAD29965.1"
/db_xref="GI:4808845"
/translation="MSSAKELKRSVLAFLVDNNPILALQVLCVCSALAVTTKLETAFTV
TLAVMEVRLSNFEVSLIRNHIPNSVRIYOMAILISLVIVVOILKAYVLYDSKQIS
TVAGLITNCIYNGRAEAFAMKSEPIPSFDIGNGIGVGFVMTVGFEPFELGSGRL
FGLFVLPISNGCWOPNGLMLLAPSAFLIGPMIAIRFRKBOYAKKE"
4027. 4623
/gene="nqre"
4027. 4623
/gene="nqre"
/note="c subunit of NOR enzyme complex"
/codon_start=1
/transl_table=11
/product="NgrE"
/protein_id="AAD29966.1"
/db_xref="GI:4808846"
/translation="MEHYISLVKSIPIENMALSFLLGMCFTPLAVSKKVTSGFLGIA
VIVLVISVPNNLVNLYLKPDALVEGLDFLNFITFGVIALVQIILEMILDRFF
PFLYNALGFLPLITYNCAIFGCVSFVWQRDYSFAESVYVYGSSGVMMALAYVALAGI
REKMKTSIDVPGLRGLGITFTTAGLMALGMSFGVOL"
```

```
gene 4657. .5883
/gene="nqrf"
4657. .5883
/gene="nqrf"
/note="beta subunit of NOR enzyme complex"
/codon_start=1
/transl_table=11
/product="NgrF"
/protein_id="AAD29967.1"
/db_xref="GI:4808847"
/translation="WSTIIIFGVMTLIIILAVLYIFAKSKLVPFGDITISINDPE
KAIYVQPGKLLTALAGVAVSVSACGGGSGCGRVKIKSGGDLIPHELHISGE
AREGERLQGVAVKADMDLEPEELFVYKMCATVVISDNKATFIKELTALPDGFSV
PERAGYIOIEAPAHNVKAYADVEDVEKRGDWKFNLEPESKVEDPIIRAYSMANP
EEFGITMVRILATPPPNPNVPGOMSSYISLAKAGDCTTISGPEEPKADTDAM
VFITGGAMAPKRSIIIFQOLKRLKSKRMSVYGAQRSAREMYVEDFDGLAENDNFV
WHCLSDPQPEDNMTGTGTGFHNVLTENTLAKHEAPDECEYTMCGPMMNAAYINMLK
NLGVEEENILLDDFEG"
BASE COUNT 1441 a 1276 c 1550 g 1649 t
ORIGIN
Query Match 38.6%; Score 518.4; DB 1; Length 5916;
Best Local Similarity 62.1%; Pred. No. 6e-120;
Matches 835; Conservative 0; Mismatches 506; Indels 3; Gaps 1;
1 atgattaaataaaaaaaggtctaactctgccatcgccgagcgagcgagcaagtcatt 60
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
34 ATGATTACAATTAANAAGGATTGGACCTTCATTCGCAAGAACTCTCCAGGTCATT 93
||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 tatgacgcccgcgcatlaccgaagtcgctgttcgacgaataatgacgcatgccc 120
||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||
94 AGTGATGTGAAGCCATCAAAAAGTCGCTTCGGCAAGATGATCGTCATCGT 153
||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||
121 cccctgatagaatacgaagaaggtgaaagccgttaaaaaaagccaaagtcgtttgaagac 180
||||| ||| ||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
154 CCTACTATGCATCCGCTAGCTAGTGAGTAATAAATCTCAGATCTTTTGAAGAT 213
||||| ||| ||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
181 aaaaagaatccggcgtagtattactgcccgcgcttcaagcaaaatcgccgatcac 240
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
214 AAGAAATCCCGGCTGTGAATTTACTTCACGTCACGCGGTAAAGTGTTGAATCAT 273
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
241 cgtgcgaagaagcgcgtactcaagtcagtcgtagtcggttgaaagcgaagaatc 300
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
274 CGTGATGCAAGCGCTCTTCAGTCTGATGATTAAGTGGCGGTGATGACCAAGTG 333
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
301 gaagtcgaagcgtacgtactgaagcgcctggcaaaatlgagcgagcaaaagtcgccc 360
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
334 ACCTTCGACAAAGTTGAAAGCCAAATTAAGCCAGTTTGAATGCGATGCAATAAAT 393
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
361 aacctgattcaatcaagcgtatgactgcgcttcgcacccgctcgttcagcaaatccct 420
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
394 CAGCTGTTGAGTCTGCTCTTGTGACTGTTTCCGTAAGTGTGCTGACCAAGGTTCA 453
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
421 gccgtagatgcccgcgttcgcaatcttcgcaaatgagtgagcaacaaatccgctgct 480
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
454 GCCATTGATTCCACTTCGAAGCCATTGTCACGCTGTGATNACCAATCCGTTGGCA 513
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
481 gccgaccccaagtcatactcaagaagcgcgcaagaacttcaaaagcggcgcttgta 540
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
514 GCTGAGCCGACCGTTGTGATGAACAAATTCGAAACCTTCGTTGCTGTGATGATGTC 573
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
541 ttagcgcgcctgacccgaagctaaatccatgtgtgtaaacgacgagcgagagctgccc 600
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
574 CTTTCGCACTGACACTGGCAAGGTAGCTTTGTA---AAAGCACACAGTTGCCA 630
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
601 tctgaaatgctgcacaatatcgaaacacatgaaattggcgccgacacacctgcgcgttg 660
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
631 CCTTTCACACAGCCGATGTTGAAGAACATGTTGACCGCCACACATCTCGACGCTTA 690
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
661 agtgcagacacatcatcttcagcgcaagtcgagcgagtaaaacgctgtgagacatc 720
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
691 GCTGACCTCACATGCAATTTCTCTATCCAGTAAGTGCACACACGTTGGCTGAGCATC 750
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||
```


TITLE	JOURNAL	source
Sellers, P., McDonald, L., Utterback, T., Fleischman, R. D., Nierman, M. C., White, O., Salzberg, S. L., Smith, H. O., Colwell, R. R., Mekalanos, J. J., Venter, J. C. and Fraser, C. M.	Submitted (14-JUN-2000)	Medical Center Dr. Rockville, MD 20850, USA
Direct Submission		
Submitted (14-JUN-2000)	The Institute for Genomic Research, 9712	
Location/Qualifiers		
1. .1382		
/organism="Vibrio cholerae"		
/strain="N15961"		
/serotype="O1"		
/db_xref="taxon:666"		
/chromosome="1"		
/note="biotype: El Tor"		
/complement(154. .1236)		
/gene="VC2287"		
/complement(154. .1236)		
/gene="VC2287"		
/note="similar to GB:U00096 SP:Q47155 PID:1552799		
PID:1786425; identified by sequence similarity; putative"		
/codon_start=1		
/transl_table=11		
/product="DNA-damage-inducible protein P"		
/protein_id="AAF95431.1"		
/db_xref="GI:9656851"		
/translation="MODIRIKIIVHDMDCFFAENVMDNPAYREIALAVGHEKRGV ISTCVQARKFCVRSAMPQAALKICPOLHVVPRMSVKSQOIQITLEORTSLIE PISLDEAVYDVSESTAYGSAATLIQAIIRDIIMQELNPAASGAVATKFLAASPLN KRDGIYVYVPRDKVOEMVDISPLEKIDGKVALERKHQGLYVAGADVRADYRKILHO ERRLASIMRKSHGIDEREVYTEREKSQVETTSQNSTOCEQVYIEQKIPYELD AALSRAHPRGIIKOGIKVFADPQDTTIEHVPALDEYDFHELLEQVLTROGREGIR LGLSLVMTKPELQMQQLSMFSPSDGMQ"		
/complement(1345. .1578)		
/gene="VC2288"		
/complement(1345. .1578)		
/gene="VC2288"		
/note="identified by Glimmer2; putative"		
/codon_start=1		
/transl_table=11		
/product="conserved hypothetical protein"		
/protein_id="AAF95432.1"		
/db_xref="GI:9656852"		
/translation="MNTPLITPGFFLAVIAAAYGIYFORKVVYKSGGGLGAVGIEKV CNCEPCKRRKREKKAARARERLAAMEKDRIA"		
/complement(1595. .2698)		
/gene="VC2289"		
/complement(1595. .2698)		
/note="similar to GB:U00096 SP:Q47155 PID:1552799		
PID:1222087 PID:1204428; identified by sequence		
similarity; putative"		
/codon_start=1		
/transl_table=11		
/product="thiamin biosynthesis lipoprotein ApbE"		
/protein_id="AAF95433.1"		
/db_xref="GI:9656853"		
/translation="MFTNMILKVSMTWFRSVSPISIFNTERSKQVRNNLVALLASL LLLACSEKPAEVOHLSGPTMGITVYIKTIQDGLADSKLTQELDRLLEEVNDQSTY RKDSICRFNOHTSESPAVSGTTLVYKAEALRLNGELGALDVTGVAIVNLMDGPE ARPDYVPEDEELNARATITGIELHTIEGNTLSKDIPELYVDLSITAKGGVGVADYD OSOGLENVETGEIRLKLGNRNDVPWRIAIEKPSVDORSVOETIIEPDYVAIVNSG YRNPEDQGVRSHTIIDPTGRIINRNVSVYLDKSCMTADGLATGLMWGEERGMA VAEANQIVLMAIVKTDGDFKEKFASSSFPELSK"		
/complement(2745. .3971)		
/gene="VC2290"		
/complement(2745. .3971)		
/note="similar to PID:893416; identified by sequence		
similarity; putative"		
/codon_start=1		
/transl_table=11		
/product="NADH:ubiquinone oxidoreductase, Na		
translocating, beta subunit"		

```
/protein_id="AAF95434.1"
/db_xref="GI:9656854"
/translation="MSTIFGVVMTLITLALVLTIFPAKSKIVPRTGITISNGDE
KAITQPGCKLTLALAGAVRVSSACGGGSCGCCRVKIKSGGDLPLPDLHISGE
AREGRILACQVAVKADMDLEPELIFGVAKWECTYISDNKATFIKELKALPDGSHV
PFRAGGYIOIEAPAHVHKYADPDEKRGDMKTNLFYKESKVDPEIIRAAVSMAYP
EEFGIIMLVNRIATPPNNPNVPCGMSYINSLAGDKCTISGPGEEFAKDTDEM
VEFGGAGAMPMRSHIIFDOLKRLSKRMSYVGAARREMEFYEDDEGLAENDNFV
MHCLSDPQPDNMTGYTGFIHNVLKENTLKHDEAEDCEYMGCPMNAAVINMLK
NLGVEENILDLDEEG"
/gene="VC2291"
/complement(4005. .4601)
/feature="VC2291"
/feature="similar to PID:663269 PID:663270; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH:ubiquinone oxidoreductase, Na
translocating, hydrophobic membrane protein NqrD"
/protein_id="AAF95435.1"
/db_xref="GI:9656855"
/translation="MEHTISLVKSIFFIENMALSPFLGMCFTLAVSKKVTSGLGIA
VIVLITISVNNLVNVLKPDALVEGDLSEFLNITFGVIALVQIILEMIDREF
PPLYNALGIFLITVNCALFGVSEFMVORDYFSAESVYGFSGVGMALAIYALAGI
REKMSVDPKRGITRIFITAGLMALGFMSPSGVOL"
/complement(4609. .5241)
/gene="VC2292"
/complement(4609. .5241)
/feature="VC2292"
/feature="similar to PID:663269 PID:663273 PID:893414;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH:ubiquinone oxidoreductase, Na
translocating, hydrophobic membrane protein NqrD"
/protein_id="AAF95436.1"
/db_xref="GI:9656856"
/translation="MSSAKELKSVLAPVLDNNPIALOVGCSALAVTTKLEAFPM
TLAVMEVTLNMFVSLIRNHIINSVRIIVOMALISLVIVDQILKAYLIDSKOLS
VPEGLITNCTIYMGRAEAFAMKSEPIPSFDIGNGIGYFVMTGVEFFELLSGKL
FELVPLISNGWYOPNGMLLAPSAFLIGFMVIMIRFKPEOVEAKE"
/complement(5241. .6014)
/gene="VC2293"
/complement(5241. .6014)
/feature="VC2293"
/feature="similar to PID:663269 PID:663272 PID:893413;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH:ubiquinone oxidoreductase, Na
translocating, gamma subunit"
/protein_id="AAF95437.1"
/db_xref="GI:9656857"
/translation="MASNNDISIKRTLEFVIALSLVCSIIIVSAAVGLRDKOKENALD
KSKLIVAGIEAKGQIYELFNKSIIEPLVDFNMGDFEGDAANADOKAAEAS
SKILTAEDOKAKIORANGVVYLVKDGKTSKVLIPDHNGLSMAYAFVAFTDGN
TVSGILTYEODGETPGISGEVENPAMRAOWMKLFDHNHPAIKIYAGGAPGSEHGV
DGLSGATLISNGVQNTFDFWGLDMGFGFPLTKVRDGLN"
/complement(6001. .7284)
/gene="VC2294"
/complement(6001. .7284)
/feature="VC2294"
/feature="similar to PID:663269 PID:663271; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH:ubiquinone oxidoreductase, Na
translocating, hydrophobic membrane protein NqrB"
/protein_id="AAF95438.1"
/db_xref="GI:9656858"
/translation="MLRQDERRVISGLKFLIEDIEHHEPFGKHEKFWALYAAAT
LFTYPLGLVTKRSSHVSDVLDKRIIMIMVLAIVPAMGMWYNAGGOAIALNHLVSGD
```

```
OLAAIVAGNMHWLTEMIGTMSDAGMSKMLGATYFLPIYATFIVGFEWELFC
MYRHEVNEGEFFVTSILFALIVPPLPLMOALGTFVGVAKREVGGRNPLDAL
AGRALEPAPYPAOISGLDWTAAADYSCATLSOAGGAGALINNAAGQITTMDAF
IGNIPGSIQEVSTLAIMTGAAPFYVMYASNRILGGWIGMILLSTFNIGSDNAM
FNMHMLHLVIGRAFGRFRRFATPDVSAFNSFGWATGILIGWCVLIRVYNAPPE
GMMALILFANLEPAPLEHDHVVERNIKRLIARYGO"
/gene="VC2295"
/complement(7254. .8642)
/feature="VC2295"
/feature="similar to PID:663269 PID:663270; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="NADH:ubiquinone oxidoreductase, Na
translocating, alpha subunit"
/protein_id="AAF95439.1"
/db_xref="GI:9656859"
/translation="WRNRIFEPPIKVVOVMITIKGDLPIAGTSPVYISDGAIRK
VALLGEEVYGMKPMHVRVGVDEVKRAQILFEDKKNQVFTSPSGVVEINGARV
LOSIVIEVAGDDOYTFDEFEANQLASLINDAIKTOLVESGLMTPRTPEFSKPAIDS
TSEALIFVAMDNPPLAEPYVYINQSEAFVAGLDVLSALTGKVVYCKKGSLSPSQ
QPNVBEHYFDGPRHAGIAGTHHFLYPSADHVANSINQDIYAVGOLFTELTYOR
VLSLGPVYVKNRFLRYTYMGASLEQVDSIEHPGKRIISGVSLSKTKGPHAYGR
YHLOVSVLREGDEKELFGWAMGKKNKFSYSTRFGLHFGQVYNMNTTNGSDRSHV
IGNREKVMPLDEPILRLDLRDLACDSDSAVRGLALELDELDALCTFVCPKREYGO
LRECLDKLEKREG"
/complement(9079. .9399)
/feature="VC2296"
/gene="VC2296"

Query Match 38.6%; Score 518.4; DB 1; Length 13382;
Best Local Similarity 62.1%; Pred. No. 6,4e-120;
Matches 835; Conservative 0; Mismatches 506; Indels 3; Gaps 1;

OY 1 atgattaaatacaaaaaggtcaaatctggccatcgcggcgagcgagcgagcaatcatt 60
DB 8594 ATGATTACATATAAAAGGATTGACCTTCATGCAAGAACTCTTCCAGGTATT 8535

OY 61 tatgacgcgcgcgcattaccgaagtcgcgtcttgccgcaagaataatgctgcgcgcgc 120
DB 8534 ACTGATGCTAAGCCATCAAAAAGTCCGCTTCTTGCGCAAGATGCTGGATCGT 8475

OY 121 ccctcgtgaaatacaagaagtgagcgcgtcaaaaaagccaaagtgcgtttgaagac 180
DB 8474 CCTACTATGATGATCCGCGTAGTGATGAATAAAAAACTCATGTTTGTGAAGAT 8415

OY 181 aaaaagaatccggcgtagtattcttcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 240
DB 8414 AAGAAAGAAATCCGGGTGTAATTTACTTACCGATCAGCGGTAAAGTGTGAATCAT 8355

OY 241 cgtgcgcaaaagcgcgtacttcagtcagtcgtagtcgcttgaaagcgaacgcaaatc 300
DB 8354 CGTGGTGCAGAAAGCGCTTCTTGTAGTGTGATTAAGGCGGCGTATGACCAAGT 8295

OY 301 gagttcgaaagcgtacgttcgaagcgtcgcaaaatgagcagcgcaaaaaagtcgcgc 360
DB 8294 ACCTTGACAAAGTTTAAAGCAATCAATTTAGCCAGTTTGATCCGATGCGATAAAACT 8235

OY 361 aactcgtatcaatcagcgttatgactgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 420
DB 8234 CAGCTGTTGAGTCTGCTGCTTGGACTGCTTCCGATCTGCTCCGTTCCAGCAAGTTTCCA 8175

OY 421 gccgtatagcgcgagcgcgttcgcacatcttcgaatgcgataagcacaatcgcgtgc 480
DB 8174 GCCATTGATTCACATCTTGAAGCATTTTGTGCTACTCTCTATGATACCAATCCGTGGCA 8115

OY 481 gccgaccctacggtcatalcaataagaagccgcgaagacattcaaacgcgcgtcttgta 540
DB 8114 GCTGACCCGACCGCTTGATCAATGAACAACTTCGTTGCTGCTGATGATG 8055

OY 541 ttgagcgcgcgtgacgcaagcgttaaatcgtgtgtgtaagcagcagcgcgcgcgcgc 600
DB 8054 CTTTCGCACTGCAACTGCGCAAGTGTACGTTGTAA--AAAAGCACCAAGCTTGCA 7998
```

OY	601	tctgaatatgctgcacatcgaacacatgaatttggcgcccgcatctgcgcgcttg	660
Db	7997	CGTTCTCAACAGCGGATGTTGAAGAACATGTCCTTTGACGGCCACACATCTTCAGGCTTA	7938
OY	661	atgtgcacgcacatcattcatcgcagccagtcggcggaataaacccgttggaccatc	720
Db	7937	GCCTGACCTACATGCAATTCCTCTATCCAGTAAGTGACGACCCAGTGGCGGAGCAATC	7878
OY	721	aattacaagacgtgatgatcgtatcgagacgttftgttgaacagcgctcgtatccgag	780
Db	7877	AACTATCAGGACGCTGATCGCATGATGTCACCTGTTCTCGACGGGAGAGCTATACCCAG	7818
OY	781	cgcgttgattgaccttggcgcgctcgtcaagtcacaacacgcgcctcttgcgtaccgtttg	840
Db	7817	CGTGTGTTTCTCTTGCCGCCCGGTGGTGACAAACCTGCTTTGGTTGCGACCTGTCAAG	7758
OY	841	ggttcgaaaggttgcttaactaccgcgcggcggaattggttgaecgcygacacccggtgatt	900
Db	7757	GGGGCATCACTAGACGAACTGGTGACGCGAAATCATGCTCTGGCAGAGTACGCATCATC	7698
OY	901	tccggttcggtatgtaacggtgcatgctgcaaaaggcgcgcatgattatgtagacgtac	960
Db	7697	TCGTGCTCTGTGCTGTGCTGTAGCAAAAGCCACTGGACACACGCTTATCTTGCGTTAC	7638
OY	961	cacaatcgaatttcggttalcgaagaagcgccgaagcaagaagctgttcgcttggttg	1020
Db	7637	CACCTACAGGCTCTGCTTCTTCGTGAAGGACGTCATTAAGCACTGTTGGTTGGCGCATG	7578
OY	1021	ccgcagccgacaataactcatccatcgacgcacacacttcggccacttccctaaacaaa	1080
Db	7577	CCGGGTAACAAACAGTTTCTGTGACTCGCTCTTTCCTTGCTACCTATTTCAAGGCCAA	7518
OY	1081	ctctcaagtlcaacgacgcgtcaacgcgcgcgcgcgcacatcgtacgtacgtacgtac	1140
Db	7517	GTGTACACATGACCCACACACACTAATGATGATGATGATGATGATGATGATGATGATG	7458
OY	1141	tatgacgcgttaatgaccttggacatccctacaccttgcgttcttggcgcatattatgc	1200
Db	7457	TATGGAAGGTCATGCTGCTATGATGAGCAACCTTGTGCTTCTGATCTGTGCT	7398
OY	1201	ggcgaatccgaacgc	1260
Db	7397	GGTGAAGTACGACGCGCGGTGCTTAGTGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGA	7338
OY	1261	ttgtgcagcttgcgttcgcccgggcaatacgaatacgaagccgcgctgttgcgcaagtcgtg	1320
Db	7337	CTGTGTACCTTGTGTTGCCAGGCAATAGAGTACGTCGTCCTGCTGAGTCTGCTTA	7278
OY	1321	gaacacattgagaagaagcgtga	1344
Db	7277	GACAGATCGAAGAAGAGGTAA	7254
RESULT 12			
LOCUS	AF165980	6048 bp	DNA linear BCT 13-NOV-2001
DEFINITION	Vibrio harveyi Na+-translocating NADH-quinone oxidoreductase		
ACCESSION	AF165980	complex operon, complete sequence.	
VERSION	AF165980.1	GI:6531420	
KEYWORDS	Vibrio harveyi.		
SOURCE	Vibrio harveyi		
ORGANISM	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.		
REFERENCE	1 (bases 1 to 6048)		
AUTHORS	Zhou,W., Bertsova,Y.V., Feng,B., Tsatsos,P., Verkhovskaya,M.L.,		
	Gennis,R.B., Bogachev,A.V. and Barquera,B.		
TITLE	Sequencing and preliminary characterization of the		
	Na+-translocating NADH: ubiquinone oxidoreductase from Vibrio		
JOURNAL	Biochemistry 38 (49), 16246-16252 (1999)		
MEDLINE	20050604		

PUBMED	10587447
REFERENCE	2 (bases 1 to 6048)
AUTHORS	Zhou,W., Feng,B., Gennis,R.B. and Barquera,B.
TITLE	Direct Submission
JOURNAL	Submitted (06-JUL-1999) Biochemistry, University of Illinois, 600
	South Mathews Avenue, Urbana, IL 61801, USA
FEATURES	Location/Qualifiers
source	1..6048
	/organism="Vibrio harveyi"
	/strain="Bb120"
	/db_xref="taxon:669"
gene	169..1509
	/gene="nqrA"
gene	169..6022
	/gene="Na+-translocating NADH-quinone oxidoreductase
	complex operon"
CDS	169..1509
	/gene="nqrA"
	/codon_start=1
	/transl_table=11
	/product="NqrA"
	/protein_id="AAFL5411.1"
	/db_xref="GI:6531421"
	/translation="MTTKGLALPIAGPSOYINDKTIKVALLGEEYVGMPTM
	VWGEQVKAQVLFEDKNPQVETPAAGKVIENKARKVLSVYIEVAGEQVFF
	DKFEAOLAGLDREVLTQLESGLWTALTRPFSVPALIESATKALFYAMDNPPLA
	AOPELLITEQOAFVAGLDLSALTEGKYVCKSGSLPSSQSNVEHEDGPHPG
	LAGTHMHFLPYNAENVVAMSINQDIARGLFLTGEIYDRVYSLAGPVNNRLLR
	TVVGSALDITDSELMPGQVRYISGVSISTQASGHAHYIGRIHQOVSVIREGDKL
	FGWATPGRKKESTISFGLHTRGQLPNNMTTNGSDRAWVPIGNERVPLDMEPTL
	LIRLDACGDTSDAQTGLALELDELDALCTFVCPGRKEYGILLRECDITVKBG"
gene	1513..2754
	/gene="nqrB"
CDS	1513..2754
	/gene="nqrB"
	/codon_start=1
	/transl_table=11
	/product="NqrB"
	/protein_id="AAFL5412.1"
	/db_xref="GI:6531422"
	/translation="MLKRFIEDIHHFEPGRKHEKWFALYEAAATLFPYRGITKRSS
	HVRSDVDIRIMTWMAFPRAMFGMVAAGCAITAIIMHVAAGDOLATVATKRRS
	LTEMIGTIGAPAGVGSMMLLATATYLPYATVYFTGEGREYVPCWRKIEVNEGFV
	TSILFALIVPPLPLMQALGIFFGVVAKEIKELFGTGRNLFAGRAFLFPAAPD
	ISGDVWNTADDFSGATLSQVNAHRSGLINNTGAPLTMDAFTGNIPGSIGEVST
	LALMGAMIVYMRITASWRIIAGVMITAVSTLEFVIGSDTPRPMNPMHNLVLGG
	FAPGFMATPDEVSAFTSGSKWYGILIGAMCVMIKRVNPAVEGMILAILPANLFA
	PLFDHVIETKNIKRLARYG"
gene	2744..3529
	/gene="nqrC"
CDS	2744..3529
	/gene="nqrC"
	/codon_start=1
	/transl_table=11
	/product="NqrC"
	/protein_id="AAFL5413.1"
	/db_xref="GI:6531423"
	/translation="WASNNDSIRKTLGVVYIGLSVCSIIIVSTAAGVLRDQXANAYLD
	KQSKIVEVAGIDAEGRKYVELFAEYIEPRLVFGKDEFEKEDGSTAANYDQRKAK
	DPAESIKRLADEDEKAKILRRANTGIYLVKNDDISKVIIPVHGNLMSMVAFFVAVE
	TDGNTVSGITVEYEGETPGLGGEVENPVRQAQFVCKFLDENHKKPAIKIIVKGAPEGSS
	EHGVDGLSGATLTGNGVGTDFPMVLGDMGCFPLAKYVRDGLN"
	3529..4161
gene	
	/gene="nqrB"
	3529..4161
CDS	
	/gene="nqrD"
	/codon_start=1
	/transl_table=11
	/product="NqrD"
	/protein_id="AAFL5414.1"
	/db_xref="GI:6531424"
	/translation="MSSAONIKHSIMAPVLDNNPIALQVLGAVCSALAVTTKLETAFVM
	TLATFVVALSNFSVSLIRNHPNSVRIITYQAAVITIASIYTVDOVYKATVYDTSRQMS

FEATURES	source
gene	10587447
gene	2 (bases 1 to 6048)
gene	Zhou, W., Feng, B., Gennis, R. B. and Barquera, B.
gene	Direct Submission
gene	Submitted (06-JUL-1999) Biochemistry, University of Illinois, 600
gene	South Mathews Avenue, Urbana, IL 61801, USA
gene	Location/Qualifiers
gene	1. 6048
gene	/organism="Vibrio harveyi"
gene	/strain="BB120"
gene	/db_xref="taxon:669"
gene	169. .1509
gene	/gene="ngRA"
gene	169. .6022
gene	/gene="Nat-translocating NADH-quinone oxidoreductase
gene	complex operon"
gene	169. .1509
gene	/gene="ngRA"
gene	/codon_start=1
gene	/transl_table=1
gene	/product="NgRA"
gene	/protein_id="AAF15411.1"
gene	/db_xref="GI:6531421"
gene	/translation="MTTKKGLALPIAGADSOYIINDKRTIKVYALLGEYVGMKPTMH
gene	VRGDEQVAKRAQVLEEDKKNPVEFTAPAAKVIIEINGAKRVLOSVIYVAGEQVDF
gene	DKEEAQALGLDREVYIKTOLPEVGLMTALRTPFSKVPALIESATKALFTVAMDNPLPA
gene	ACEPLILITEQOEAFFVAGLDILSALTECKYVVCSGTSLPSSOSNVEHEDGHPHGF
gene	LACTHHPLFIPVAENVMASINODVYAFGLRLTGLTYLDVRYSLAGPVNNPRLR
gene	TYVAGLDELTDSELMPEGEYRVISGVSITQASGRPAIVIGRIHOQVSVLREGDKEL
gene	FGNATFGKNKFSITKSFILGIFKQLEPMTTNGSDRANVPIGNETVRMLDEPPL
gene	LEHDLACGDTDSQTLGALELDEDEALCTCPGKYEYQLRECDITVKEG"
gene	1513. .2754
gene	/gene="ngRB"
gene	1513. .2754
gene	/gene="ngRB"
gene	/codon_start=1
gene	/transl_table=1
gene	/product="NgRB"
gene	/protein_id="AAF15412.1"
gene	/db_xref="GI:6531422"
gene	/translation="MLKKFLEDEIHHPEPGKHEKMPALYEAAATLTPFGIVTKRSS
gene	HYRDSYDLKRIIMVAFVPEPMKGMVAGCOALAINMVGADOLATYAGNMNHW
gene	LTLEMGGTIGAEAGVSSMLLGATFTPLPYATVFTYGGREYVLCMKRIEVEGFPV
gene	TSLDLVLIPTPLTQALDGLTFEVVAAKEIGTGRNPLNALGRALFLFAPPAQ
gene	ISGDVVAADGFSGATLSQWMAHRSGLINNITAPILMDAFTGNIDGSIGEST
gene	FLMGMAAIYVMRIASWRIAGVMIGMIASTLPVNIIGSDTIPDMFNMPHMHLVLGG
gene	LALGMPFAMTDEPASAFSTNSGKMWYGLILGAMCVMIHVNPAYPEGMILALPANI
gene	FLPDHVIKKNIKRRRLARYKG"
gene	2744. .3529
gene	/gene="ngRC"
gene	2744. .3529
gene	/gene="ngRC"
gene	/codon_start=1
gene	/transl_table=1
gene	/product="NgRC"
gene	/protein_id="AAF15413.1"
gene	/db_xref="GI:6531423"
gene	/translation="MASNNDISIKRTGLGVYIGLSLVSIIYVTAAGLRDQOKANAYLD
gene	KOSKIVAGIDAEGRKPELFAEYIEPLVYFKTDFEKAEDSGSTAAVYDORAKAK
gene	DEASQILTADEDAKILIRRAVNTGIYLVKNGDIDISKVILPYVNGSLMSMYLFVVE
gene	TGNSVYSGITTYECCGFTPLGGEVENPVVRAQVYKKFLDENHKPAIKIYKGGAPGS
gene	EHGVQIGSATLTGNNGVCTPFDWLGDMKGFPLAKYVRDGLN"
gene	3529. .4161
gene	/gene="ngRD"
gene	3529. .4161
gene	/gene="ngRD"
gene	/codon_start=1
gene	/transl_table=1
gene	/product="NgRD"
gene	/protein_id="AAF15414.1"
gene	/db_xref="GI:6531424"
gene	/translation="MSSAONIKKRSIMAPVLDNNPILAOYLGVCSALAVTNTKLETAFM
gene	TLAAFEVRLSNFSVSLINNHIPNSVRIIVQMAILASITVYDOVYKATVYIKTSQAS

VPVGLITNCIVMGRAEAFAMKSAVPVSLIDINGLNGLYGFLVITGVFPFRELFGSKL
FGMEVPLIVSMGWKQVPGMLILAPSAFLLIGFLIIVIRILKPEOVEAKE"
gene
4169..4765
/gene="ngre"
4169..4765
/gene="ngre"
/codon_start=1
/transl_table=1
/product="ngre"
/protein_id="AAF15415.1"
/db_xref="GI:5631425"
/translation="MEHYISLVKSIFFIENMALSPFLGMCPTLAVSKKVKTSFGLGVA
VVVVLTAVPNNLVNLYLKENALVEGVDSFLNFIPIFYIAALVQJLIEMLIDRF
PLYNALGIFLPLITVNCARFAGVSEFMORDPAFESIYGHGSGVMALAIYALAGI
REKKKYSVPPGLKGLGITFTITVIGMALGPNFSVQL"
gene
4799..6022
/gene="ngre"
4799..6022
/gene="ngre"
/codon_start=1
/transl_table=1
/product="ngre"
/protein_id="AAF15416.1"
/db_xref="GI:5631426"
/translation="MDILGVVFTLIVLALVILVIFAKSKLVPDGIITISVNDPSL
ALVOPGKLISALAGVSVSACGGSGGSCOCRRKIKSGGDLPTLELHDITKGBA
REGERLACQVAMKMDIPEIRFVYKMECTVVISNDKATPIREKLKIOIPGESVP
FRAGYIOLIEAPAHVKYADVIDPEYREDMEKFNLPRESKYNETIRAYSANNP
EFGIIMLVRIATPPNNPDVAFLINSFISLKBEDKCTISDPEDGALADNDENFV
FVGCGAGMAPMRSHTFDOLRLHSKRKMSFWGARSKRMEFYEDDIFGLAADNENFV
HCALSDPLPEDNWDGYTGFIVHVLVYENYLRDHEAPRDECYMCGPMMNAVYIGMLKD
LGVDENTLIDDFGG"
BASE COUNT 1571 a 1351 c 1455 g 1671 t
ORIGIN

Query Match 38.2%; Score 513.6; DB 1; Length 6048;
Best Local Similarity 61.9%; Pred. No. 9.8e-119;
Matches 832; Conservative 0; Mismatches 509; Indels 3; Gaps 1;

QY 1 atgattaaatacaaaaaaggtcctaactcgtccatcgcgagcgagcgagcaagtcatt 60
DB 169 ATGATTACATTAAGAAAGGCTTGCTCTTCTATCGCAGAGCTCCATCCAGTGATT 228
QY 61 tatgaagccggcccatcaccgaagtcggttgcttggcgaaagataatgctgcgattgcg 120
DB 229 AATGATGTAAGACCATCAAAAAGTCCTTGGCGAAGATACGTTGGCATGCGT 288
QY 121 ccctcgatgaaatacaaggaagtgaaagccgtcaaaaaagccgaagtgtgttgaagac 180
DB 289 CCTACCATGTCATCTCGCGTTCGCGATGAGTGAAGTGAAGCGCAAGTTCTTTGAAGAC 348
QY 181 aaaaagaatccgggtagattactactgcgcggttcgagcaaatccgcgtatttac 240
DB 349 AAAAAAGAACCTGGCGTTAAGTTCACGCTCCAGCACCGGTAAGATGAAATTAAAC 408
QY 241 cgtgggaaagaagcgctactactcagtcagtcgtgattgcgttgaagcaagaagaaac 300
DB 409 CGTGGCGCTTAAGCGGTACTTCAATCAGTTGTGATTAAGTAGAGGGAAGACAGGTT 468
QY 301 gaatcgaaacgctacgtaccgtgaagcgctgcaaaattgagcagcaaaaagtgcgcgc 360
DB 469 ACCTTTGATTAAGTTCAAGACAGCTCAACTAGCAGGTCTAAGTCGCAAGATCAAGACT 528
QY 361 aacctgaatcaatcaggtttagactgccttcgcaaccggtccgttcagcaaaatccct 420
DB 529 CAATTGTGTAATCTGGCCTTTGGACCGCTTACGTACTCTCGTTCAGCAAGGTTTCCA 588
QY 421 ggcgttagtcgagcgcgcttcgcacatctcgtcaatgcgatgcgaacgaacgctgct 480
DB 589 GCATATGAGTCTCTACTTAAGGCTATTCTTGTAACTGCGATGATACCTAATCCACTAGA 648
QY 481 gccgaacctacggttcacatcaaaagcgcgcgaagaactcaaacgcgctgttgta 540

DB 649 GCACAGCCTGATGATGATCATCATGMAACAAGAGACTTCGTGCAAGCTTAGACATT 708
QY 541 ttgaagccgcttgaccgaacgctaaatccatgtgtttaagcgagcgagcgagctgccg 600
DB 709 CTTTCAGCCCTGACAGAGGAGGCTTACGATGTAA---CTTGGCACCTAGCTTGGCA 765
QY 601 tctgaatctgcgcacaatccgaatcgaatctgagcgccgcgacatcctgcggtctg 660
DB 766 AGCTCTTCTAGTCAAGAGTGAAGACACGCTTCGATGATGTCACACCGACAGGCTTT 825
QY 661 agtgcacgcaatcattcattcgaagccagtcgagcgagcaataaaacgctggaacatc 720
DB 826 GCTGGACCCACATCATCTTCTATACCAAGTAATGCGAAATGCTGGCGTAGACATC 885
QY 721 aattacaagcgtgattgtgtatcgtgaggttgcttgcgaacgagcgcttgaataacgag 780
DB 886 AACTACCAAGAGCTTATGCGCTTGCGTACGCTTTTCTAACGGGTAGCTTTACCTGAT 945
QY 781 cgcgtgttccttgagcgccctgcgaagtcacaacaaacgcgcctcttgcttacggtttg 840
DB 946 CGTGTGTGTCTCTAGCTGCTGCTGTTGTGAACAACCGCCTTCTGCTACAGTTGTA 1005
QY 841 ggtgcgaagtgctcaactaccgcgcgagattggttgaagcgagacaacgcggtgatt 900
DB 1006 GGTGCTTCTTGAAGACCTTAACCTGACACGAGTTGATGCTTGTAAGTTCTGTGATT 1065
QY 901 tccggttgatgattgaagcgctgcgattgcacaagcgcgcaatgatttatttgaagctac 960
DB 1066 TCAGGCTCGGTACTTCTTGCTACTCAAGGCTCTGCTGCACATGCTTACAGTCTGCTTAC 1125
QY 961 cacaatcagattccgttactcgaagaagcgcgagcaaaagcgtgttcggtcggttgcg 1020
DB 1126 CACCAACAAGTTCTGTGTGTACGCCAAGGCGHGAACAAGACTGTTGGCTGGCTTACA 1185
QY 1021 ccgcagccggacaatactccatcagcgacacactctcggtccatttcccaaaaaa 1080
DB 1186 CCTGTGTAAMAACAAGTCTCAATCAAGAAATCTTCTTGCTGTCACATCTTCAAGGTCAG 1245
QY 1081 ccttcgaattcagcagcgcttcaacgcgcgagcgcgccatcgttgaacgagtcgacat 1140
DB 1246 TTGTTCAACATGACAAACAAACGACCAACGATGACCGTGCATGTTGCCAATCGCAAC 1305
QY 1141 tatgagcgtaatgcggttgcgttcgacatccctgctactcgttcttgcgagattaatgctc 1200
DB 1306 TACGAGCGGTAATGCCGCTGTGATGGAACCGACCTTGGCTACTTGCATGATATGCA 1365
QY 1201 ggcgataccgaacgcgcgagcgttgggttgcgttgaattggaacgaagaagcctgc 1260
DB 1366 GCGGATACAGACAGTCTCAAAACACTGGGTCGCTTGAACTGATGAAGATCTAGCA 1425
QY 1261 ttgttagcgttgcgttcgagcggaataatagatagcgccgctggtgcgaagtgctg 1320
DB 1426 TTGTGTACTTGTGTATGTCCTGTTAATATGAGATGCGTACGTAAGTCTGTAATGCTTA 1485
QY 1321 gaaacattgagaagaagcgctga 1344
DB 1486 GATPACCATGCTGAAGAGAGGCTAA 1509

RESULT 13
VANORBOL 5078 bp DNA linear BCT 15-JUN-1999
LOCUS V.alginolyticus bola, ngfA, ngfB, ngfC, ngfD and ngfE (partial)
DEFINITION genes.
ACCESSION Z37111.1 GI:563268
VERSION 237111.1
KEYWORDS NADH:ubiquinone oxidoreductase subunit A; NADH:ubiquinone oxidoreductase subunit B; ngfA gene; ngfB gene; ngfC gene; ngfD gene; ngfE gene.
SOURCE Vibrio alginolyticus.
ORGANISM

REFERENCE AUTHORS TITLE	JOURNAL MEDLINE REFERENCE AUTHORS TITLE	FEATURES source
Bacteria; Proteobacteria; gamma subdivision: Vibrionaceae: Vibrio. 1 (bases 1 to 2384) Beattie, P., Tan, K., Bourne, R. M., Leach, D., Rich, P. R. and Ward, F. B. Cloning and sequencing of four structural genes for the Na(+)-translocating NADH-ubiquinone oxidoreductase of Vibrio alginolyticus FEBS Lett. 356 (2-3), 333-338 (1994)	Ward, F. B. Direct Submission Submitted (09-SEP-1994) for Cell and Molecular Biology, Mayfield Rd, Edinburgh, Scotland, EH9 3QR	location/Qualifiers 1..5078 /organism="Vibrio alginolyticus" /db_xref="taxon:663" /clone="10-2" /clone_1lb="Lambda library" 26..334 /gene="bola" 26..334 /gene="bola" /function="Cell division" /codon_start=1 /transl_table=11 /product="Bola" /protein_id="CAAB5475.1" /db_xref="GI:663269" /db_xref="SWISS-PROT:Q65585" /translation="MIQETLEKLSQPSYTKVNIENSYMHNPQSGSESHFKYIVVS DSRAGLIGRHRQVHQIADLDELNHIHALIHITYDEEWKREDDGAPDSFMCGGGH "
808..2148 /gene="nqrA" 808..2148 /gene="nqrA" /codon_start=1 /transl_table=11 /evidence=experimental /product="NADH:ubiquinone oxidoreductase subunit A" /protein_id="CAAB5476.1" /db_xref="GI:663270" /db_xref="SPRMBL:O56586" /translation="MITIKKGDLPIAGPFSOVINDKTIKVALIGEERYVAKRPTMH VRGDSVKAQVLFEDKKNPKGFTFPAAGKVLEVNKGAKEVQSVIEVAGEQVTF DKFEAOLSGLDREVIKTQLVDSGLWTALTRPFSKVALESSTKAIFVAMPTNPLA APPELLINQOEAFIAGLDILSLTEGKVVYVCSGLPSSQSNENHVFQDPHPAG LQTHHFTLPVNAENVAMSINVQDIARGLKLTGELYTDKRVSLAGVYVNNPRLYR TVIGASLDLTDNELMPGEVRVYSGSVLGTNAHTGPHATIGRTHQOVSVIREKEKL FQWAMGKPKFSYTRSFGLGHVRKGQLFNMTTNGSDRSKVPITGNERVPLDMEPTL LHLDAGDPTDSQAQALGALDELDELALCTFVCPKRYEYGTLLRECLDTLEKRG"	gene CDS	gene CDS
2118..3398 /gene="nqrB" 2118..3398 /gene="nqrB" /codon_start=1 /transl_table=11 /product="NADH:ubiquinone oxidoreductase subunit B" /protein_id="CAAB5477.1" /db_xref="GI:663271" /db_xref="SPRMBL:O56587" /translation="MPRYRREGRVIFIMALKLEDEIHHPREGGKHEKMPALYEAVAT VPTPGVITNKSXVSDVLEKRIIMTWFAVFPAPFWMYNAAGAIALAHMYAGD QIATVYISGNMHWLTLEMGITLAADAGVSKMLGATVFLPIATVFLVGGFEVFLC AKRKEHVENGFVTSILFALIVPTPLIMQAALGTFGVVAKEIFGGGGRFLNAL AGRALFPAYPALQISGDVYTWAVDGSFATLSOMAGNGALVNVITQSPITPMNLF IGNIGSISGEVSTLALIMAAIIVYRIASIGRTIIGVMIGMIAVSLTFVNTGSDTNPQ FMPPHMHLYLIGCFAPGMEFMAITDPYSASFYTKGKRWYGITLIGAMCVIRVNPAYPE GMMLALIRANLFAPIFDHVIETKNIKRRILARYK"	gene CDS	gene CDS
3388..4158 /gene="nqrC"		

CDS	3388..4158	/gene="nqrC" /codon_start=1 /transl_table=11 evidence=experimental /product="NADH:ubiquinone oxidoreductase" /protein_id="CA85478.1" /db_xref="GI:663272" /db_xref="SPTREMBL:Q56582" /translation="MASNDSIKRTLGVIIGLSVCSIIYSTAAVLGRDKORANAVLLIKSLIVEAGIDANGKVEPLEFAEYIEPLVDJETGNTFEGMASTYDQREASKDAERIALTPEEDVADIRRRANTATVAIVLVKDDEVOKYLPHMGAGIMSMYAFVAVETDGNNVSALPIYRDOGERTPGAGEVENPSRPDOFGIKLYINEDHQAIAIKVKGAPOGSEHGVIQLSGALTISNGYQHTEFDWFLGDGKFQEPFLAKVRDELN"
gene	4158..4790	/gene="nqrD" /protein_id="CA85479.1" /db_xref="GI:663273" /db_xref="SPTREMBL:Q57095" /translation="MSSAONKKRSILAPVLDNNPILAOVLGVCSALAATTKLETAFFNVLTAEFTVALSNFSVSLIRNHIPNSVRIIVOMAILSVLIVDVOLKAAYDISKOLSVEFGIITNCINGRAEAFAEMKSAPEPSLIDGNGIGYGFEVILTIVGFPRFLPSGKRIFGLGVLPVINGNWGYQPNNMLTLAPSAFPLIGLIIVIRLKPEYPAKE"
CDS	4798..5078	/gene="nqrE" /protein_id="CA85480.1" /db_xref="GI:663274" /db_xref="SWISS-PROT:O56589" /translation="MEHYISLWKSISSSHALSFFLGMTFLAVSKRKVTSGFGLVAMVVVITIAVPNNLVNLVIRENALVEGSDFLFNFITGIVIALVOIL"
BASE COUNT	1366 a 1068 c 1207 g 1437 t	
ORIGIN		
Query Match	36.3%;	Score 488; DB 1; Length 5078;
Best Local Similarity	60.7%;	Pred. No. 2.8e-112;
Matches 816; Conservative 0;	Mismatches 525;	Indels 3; Gaps 1;
QY	1	atgattaatacaaaaaggtcctaatactgcgccatcgccggacggcggaagtcatt 60
Dd	808	ATGTATTACATAAAGAAGGGCTTGSGATCTTCCTATTCGACGAATCCATCCCAGGTGATT 867
QY	61	tatgacggcccccatcacccaagtgcggttgcgttgcttgcggaaatatatgctgcgatgcg 120
Dd	868	AATGATGCGTAAGACCATTCAAAAAGTCGCCGTTGCCTTGCGAAGAGTAGTGTGGCATCGCT 927
QY	121	ccctcgatgtaaatcaagaaggtgtgaagccgtfcaaaaaaggccaagtgcgtttgtaaagac 180
Dd	928	CCTAACATGCGATGTCGCCGTTGCGTGATGAAGTGAAGAAAAAGCCGACAAGTCTTTTTGAAGAT 987
QY	181	aaaaagaatccggcgctgtagttaactgcgcggcttcagaccaaatgcgcgttatccac 240
Dd	988	AAAAAGAACCCCTGGCGTAATTTTACTCTCACACGACGACCGGTAAAGTGATCGAAGTTAAC 1047
QY	241	cgtggcgaaaaagcgcgtaactcaagtcgtagtgcgcgtttgaagcaaacgaaatc 300
Dd	1048	CGTGGCCCTAAACGTCCTTCCTCAATCTGTAGATTTGAAGTGGCAGTGACGAGCAGCGTG 1107
QY	301	gaagtcgaagcgtacgtaccctgaagcggtgtgcaaaattgagcagcgaaaaagrtgcgcgc 360
Dd	1108	ACATTTGATTAAGTTTGAAGCCGCTCAACTTTCAGAGGTCTTAGATGCTGAAGTGATCAAGACT 1167
QY	361	aacctgatacataagcttatggactgcgttcgcagcccgctccgttcaagcaaatccct 420

/db_xref="GI:2558473"
 /translation="MTIKKGLDPLTAPGSPSYINDCKTIKKVALLGEETVGMREPTM
 VRVGEVKKRQVLFEDKKNGVATPAPAKVLEVNNGARKLVDSVILEVAGEQVETM
 DKPELAQSDLEREVLTQLVDLSGLMTALTITRFPSSKVALSESTKALFVTAMDNPIL
 AKPELILEQDEAPFIALGDIILSALTEGVKVCSSGLPRSSQSNVEHEVVDGHPAPA
 LAGTMMHIEFVPAENRVAWSINQDVIAFGKILFTGLYDTRVSLAGPVVNNPPLV
 TVIGASLDLITDNEMLPGEVRVLSVLTCTHATGPAAYIGRYHQOYSVLRGEKEE
 FGWAPRKKGFVTSRSLFGLHVEFGGLFNMTTTNGSDRSWVPICGNRYVMDLDEPTT
 LLRDLCAGDTSDAQAQLGALDEBDLALCTFVCPGKIEYGLLRECI DTIEKG"
 1205
 /gene="nqr1"
 /citation=[2]
 /replace="c"
 1234
 /gene="nqr1"
 /citation=[2]
 /replace="g"
 1256
 /gene="nqr1"
 /citation=[2]
 /replace="t"
 1732..2976
 /gene="nqr2"
 1732..2976
 /gene="nqr2"
 /note="Hydrophobic membrane protein with pI 8.14; nqrB,
 Nqr2 subunit of Na-NQR complex"
 /codon_start=1
 /transl_table=11
 /evidence=experimental
 /product="Nqr2 subunit of Na-translocating NADH-quinone
 reductase complex"
 /protein_id="BAA22911.1"
 /db_xref="GI:2558474"
 /translation="MALKFLEDIEHHFEPGKHEKMFALYEAAIVFYFNPGLYTNKS
 SHVDSYDLKRLIMVWFVAPFPMFGMGNAGSQAIALNHWYAGSOLAIVISNMHRY
 WLTMDVGTITADAGVSGKMLGATFYELDIYATVELVGVGHWELFCVRKHEHNEGFF
 VTSILFALIVPTPLPMLQALCITTPGVYAKELFGTGKFLPLPAGRAFLFEPAYPA
 QIISDVGMTADHGSQATASQAOGSNGALVNTVGSPIITMDARIGNIPGSGEVS
 TLALMIGVAMIVMRIAASWRIIAGVIGMIAVSTLENVIGSDINPMFMPEMHHLLVIG
 GFAPGEMFATDPVASFTNKGKMWYGLIGACWMI RVNPNAYPEGMLALIFANLF
 APLDHVIEINIKRIRIARYCK"
 181
 /gene="nqr2"
 /citation=[2]
 /replace="c"
 2400
 /gene="nqr2"
 /citation=[2]
 /replace="t"
 2417..2418
 /gene="nqr2"
 /citation=[2]
 /replace="tc"
 2614
 /gene="nqr2"
 /citation=[2]
 /replace="g"
 2966..3736
 /gene="nqr3"
 2966..3736
 /gene="nqr3"
 /note="Acidic protein with pI 4.83; nqrC, Nqr3 subunit of
 Na-NQR complex gamma-subunit"
 /codon_start=1
 /transl_table=11
 /evidence=experimental
 /product="Nqr3 subunit of Na-translocating NADH-quinone
 reductase complex gamma-subunit"
 /protein_id="BAA22912.1"
 /db_xref="GI:2558475"
 /translation="MASNNDISIKKTLGVVIGLSLVCSIIIVSTAAGLRDCKRANAYLD
 KQSLIEVAIDANGKRVPELFAEYIEPRVLDLEFGNFTEGNASTYDREASKDAERS

```

conflict
/gene="nqr3"
/citation={2}
/replacement="c"
3514. .3515
/gene="nqr3"
/citation={2}
/replacement="ca"
3736. .4368
/gene="nqr4"
3736. .4368
/feature="nqr4"
/feature="hydrophobic membrane protein with pI 9.17; nqr4,
Nqr4 subunit of Na-NQR complex"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="Nqr4 subunit of Na-translocating NADH-quinone
reductase complex"
/protein_id="BAA22913.1"
/db_xref="GI:2558476"
/transl_table="MSSAONVKKSLIAPYLDNNPALLQYLVCSSALAVTKLEAFYWM
TLATFVATLSNFSVSLIRNHIPNSVRIIVQMAIIASLIYDQVLKAYLDISKQSL
VFQGLITLNCIVNGRAEAFMKSAPVSLIDIGNLGQFVLIYGFPELFFSGKRI
FGLEVLIVNSGVNGVQPNGLMLLAPSLIFGLFWIRILRKEQYEAK"
4376. .4972
/gene="nqr5"
4376. .4972
/feature="nqr5"
/feature="hydrophobic membrane protein with pI 7.00; nqr5,
Nqr5 subunit of Na-NQR complex"
/codon_start=1
/transl_table=11
/evidence=experimental
/product="Nqr5 subunit of Na-translocating NADH-quinone
reductase complex"
/protein_id="BAA22914.1"
/db_xref="GI:2558477"
/transl_table="MEHYISLVKSIFFENMLSEFLGCMFLAVSKRYTSFGLGVA
VVVLLTAVPNNLVNVLRENALVEGVDLSFNEIFPIGVALLVQLEMLDREF
PPLYNALGFLPIITVNCATIRGVSAFPMQRODVNFESIIYGGSSGVGMALAIYALAGI
REMKKSYDVPGLRGLGITFTTIGLMLGPMFSGVQL"
4410
/gene="nqr5"
/citation={2}
/replacement="n"
4419
/gene="nqr5"

conflict

conflict

Query Match      36.0%; Score 483.2; DB 1; Length 6908;
Best Local Similarity 60.5%; Pred. No. 4,66-111;
Matches 813; Conservative 0; Mismatches 528; Indels 3; Caps 1;

1 atgattaaatcaaaaaaaggtctaattctgcacatcgcgcgagaccgagcaagtcatt 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
386 atgatttacaattaaagaaagggcttgatcttccattgcgcagaaactccatcccggtcatt 445
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

61 talagagcccgccatattaccgaagtcgctgtcttgcggaagaatatgcgcattcgc 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
446 aatgatgtgtaagaccatcaaaaaagtgcccttgctttggcggaagatgcttgscattccgt 505
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

121 cctcgatgaaatcaagaagtgtaagcgctcaaaaaagcgaagtgcgtttgaagac 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
506 cctaccatgcattcccgcttggtgtagtgaamaagcaaaagcccaagttctttttgaagat 565
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

181 aaaaagaatccggcgctagatattcttcgcccgcgttcgaagaaatcgccgctattcac 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
566 aaaaaaacccttcggcggaattttactgcaccacgacgcgcggtaaaagtctccgaagttaac 625
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

241 cgtgcgaaaaagcgcgctacttcagtcagtcgttgatccgcttgtaagcaacgacgaatc 300

```



```

SVALKFEFEREGYNITPEDEGDEILRNGTVDLGFSYMASTVKSQDKNDNTGDIYVNG
LPNGVENPYKITSDDMGMAIDPTGLRTYLNFRFYDYOPLFTVENGFGADVLEEDGKY
HDPRELOYKXSHLEALKKAVTYDGVLDIGTYPMSIIDIVSTTGEKMKRYGMITYDRD
NEGNGSMKRYKXDSFEMYNVIOFTNGEEL"
4733..5449
/gene="yydk"
4733..5449
/gene="yydk"
/codon_start=1
/transl_table=11
/protein_id="BA11271.1"
/db_xref="GI:1064785"
/translation="MKMLKYQIATEIEITYIEEHQLOQGRKLPVLETLMQFEVSKST
ITKSLLEOKGALFQVNGSGIFVRKHKRGYISLNSGCFKLEDFPNVTSKYELD
VRRPPEAENLNGMEDIEDIYYKRYRINRGQITLCEESYTKSIVTYLNEIYSHSI
FHYIRGELGKIFGSDLPFHVQOLNEEAAYLGLLEAGLPRTYIESIRHLTNGQFEDIS
KISYNYEOSOFVQVNSFLL"
complement(5766..6488)
/gene="yydu"
complement(5766..6488)
/gene="yydu"
/codon_start=1
/transl_table=11
/protein_id="BA11272.1"
/db_xref="GI:1064786"
/translation="MKLEPKKSIKSVIIITLIGAMEVFLIGYFLLGIDKSVNTPTE
MEFFSYVATQFGLMESEVLAIFNREYSKNMILFYKLIGENIYFFPKKIVLFL
ECFAFITGLIISLMDHDFSHFALLFLFSAVILOYIILIGISVCPNLLISIGS
IVVMSTVILVAISNKTGFIAPFEAGNTMYPRIERVLQSDNMNIGNDVLFILLYV
SIITINAVLRPSKRWTKMGL"
complement(6509..7138)
/gene="yydi"
complement(6509..7138)
/gene="yydi"
/codon_start=1
/transl_table=11
/protein_id="BA11273.1"
/db_xref="GI:1064787"
/translation="MNINAYTLKVKGKTLLODTDLMESSGKIHVVYKNGKVGKSQLAK
DFLNSKRIRDRIONVSLISSSNIPDVGSKDFLHFLSKFEDAKMDIKAYILML
DNIDGKYLKINISDQOKOKLISFLLEDKNILVIDEINSLDKKVIETIEHGLNKKI
QENPEKIIINTHTSDKATIEGDYIIFNHQEQHSHVSKLEIVINE"
complement(7288..8046)
/gene="yydh"
complement(7288..8046)
/gene="yydh"
/codon_start=1
/transl_table=11
/protein_id="BA11274.1"
/db_xref="GI:1064788"
/translation="MKILIKYEDGKEYVLVONNVFKDKKSGEYKNSLNSLSDKQLLR
FKMKREKVSPEFFYLFUSFTALMTLNYIHLKILONGLSVFGKMKWIIIVYFINN
IVLHELGHISLKEFGKNFVGVKLNLYVPAPVQVLETNYLSRNEKIIIVHGLGF
INVILLINLEINOFTESSEALTMAFMLFSSILLNMLIPILNSDGKILLAPLSDEY
SRKTNHMLVYITIOIGLAVNSVVMHILYLVN"
complement(8027..8986)
/gene="yyg6"
complement(8027..8986)
/gene="yyg6"
/codon_start=1
/transl_table=11
/protein_id="BA11275.1"
/db_xref="GI:1064789"
/translation="MYNKTVSINLDSRCNASCDCFCSSSPSTTRMEKEYIRELVTB
FANKKIQVISTFGGEVFLDYKFLKELMEIILPYBKQITLISNGFWSLKKVOEYEH
DNMSLNVIALTISYDEYHAPVKSISIKNILEHSKRYPDIDISLMNAVTKRQMSHIL
BELGDSIIIGVKTREPMISVGAKTRIKQENIHKRYSLDEDSLHCPCGDIYHHRGE
IYCCSPATREKTLTREBNQSEERYEAKNSNLLTLKREGKRWFLNITLKENNKI
BEEDIPYESSIGVCGSLFNSAEKINIFYMEKYENENFV"
complement(9044..9193)

```

Query Match 3.2% Score 42.4; DB 1; Length 36200;
Best Local Similarity 64.0% Pred. No. 8.9;

```

Matches 64; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 77 ttaccgaagtcgcgttgcgttgcgaagaataatgctgcagtcgcccctgcagtaanaatca 136
4337 TTGGCAGATCGGGCTGGATACGGTACGACATTGACAGCGACGCCCTTTCTTCTGAAGTAA 4278
QY 137 aggaagtgtaagccgtcaaaaagcccaagtcgttttga 176
4277 AAGAAGCGATACCGTCAAAAAGCGAAGTGGCTTTGTA 4238
RESULT 21
AP003456 151329 bp DNA linear HTG 29-MAR-2001
LOCUS Oryza sativa chromosome 6 clone P0542E10, ** SEQUENCING IN
DEFINITION PROGRESS ***, in ordered pieces.
ACCESSION AP003456 GI:13486925
VERSION AP003456
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Oryza sativa (cultivar:Nilpobare) DNA, clone:P0542E10.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (sites)
REFERENCE Sasaki,T., Matsumoto,T. and Yamamoto,K.
AUTHORS Oryza sativa nilpobare(GA3) genomic DNA, chromosome 6, PAC
TITLE clone:P0542E10
JOURNAL Published Only in Database (2001) In press
2 (bases 1 to 151329)
3 Sasaki,T., Matsumoto,T. and Yamamoto,K.
4 Submitted (28-MAR-2001) Takuji Sasaki, National Institute of
5 Agrobiological Resources, Rice Genome Research Program; Kannonnai
6 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
7 (E-mail:tsasaki@agr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
8 Tel:81-298-38-7441, Fax:81-298-38-7468)
9 NOTE: It currently consists of 1 contigs. Gaps between the contigs
10 are represented as runs of N. The order of the pieces is believed
11 to be correct as given, however the sizes of the gaps between them
12 are based on estimates that have provided by the submitter. This
13 sequence will be replaced by the finished sequence as soon as it is
14 available and the accession number will be preserved.
15 * NOTE: This is a 'working draft' sequence.
16 * This sequence will be replaced
17 * by the finished sequence as soon as it is available and
18 * the accession number will be preserved.
19 location/Qualifiers
20 1.151329
21 /organism="Oryza sativa"
22 /cultivar="Nilpobare"
23 /db_xref="taxon:4530"
24 /chromosome="6"
25 /clone="P0542E10"
BASE COUNT 40598 a 33769 c 34863 g 41449 t 650 others
ORIGIN
Query Match 3.2% Score 42.4; DB 2; Length 151329;
Best Local Similarity 46.1% Pred. No. 9.9; 166; Indels 0; Gaps 0;
Matches 142; Conservative 0; Mismatches 166; Indels 0; Gaps 0;
QY 1036 tactcatacagcgaccacactctgcgcattctcctaaacaactctcaagtcacg 1095
25278 TCCCATCAGCGTTTCGCCCTCGCGCCCTCTTTCAACTCACTTTCAGCGGCAATC 25337
QY 1096 acaagcgtaacagcgagcgacgcgcatgtacacgacgtgcacattatgagcgcgtaatg 1155
25338 TCCGACGTCGCGGCATCTCGACCCCAAGGTCGGGTATCTCTCCAAAGCGGCTCTCG 25397
QY 1156 ccgttgagacatccgctcactctgttgcgcgatttaatcgtcgcgataacgacagc 1215
25398 CTCTTCCTCCCGTCATGCTTCTTCGAGAGCGCAAGAACGCCGCTGCGCGGCC 25457

```



```
CDS      /gene="yxjB"
          complement(3797..4645)
          /gene="yxjB"
          /function="unknown"
          /note="similar to hypothetical proteins"
          /codon_start=1
          /transl_table=11
          /protein_id="CAB15937.1"
          /db_xref="GI:2636447"
          /db_xref="SWISS-PROT:P42313"
          /translation="MKRTVDFSMFRCPIDSSMDASGKSLICTERCHTPELDSRGV
          NFKTPKTSVGAELFEARSRLIGEGCFDPDLDAIAELISHPKSGHEATLIDSGG
          BESHINALCGEPYAGKAATGIDIDSKDILKASKAFKDLMAVADVARAPFHDROP
          VVLSIFSPSNVAEFHRLKNDMLIKVPRSDYLIELRQFLYDTSPPRTYSNTAAVER
          FTANAASHRPVRLRYVKTLDDQAIIHMLKMTPLAMSAKPDVSLKEMKSADITVDV
          ILICMK"
          5261..6489
          /gene="yxjA"
          5261..6434
          /gene="yxjA"
          /function="unknown"
          /note="similar to pyrimidine nucleoside transport"
          /codon_start=1
          /transl_table=11
          /protein_id="CAB15938.1"
          /db_xref="GI:2636448"
          /db_xref="SWISS-PROT:P42312"
          /translation="MYFLNLVGLIVMAVFLCSPQKKIKWRPITLLIVLELLITM
          FMLGTGVSGMAIGKIDPEPTLICASDGIARAAPSVMNEYDEFFSALPIPIFYLV
          FFDITLYFGIILPIDIKGWISKASRLPKLSFPTSIOMFLGNLTALAVTROULTVL
          SNNRLITGLMSMSISGSIISGLSVAPATVFAIPLNCNALIILANLINVAVE
          DEDILYTPKREKDKFSTISNSMLVGMNVIVIIAMVYVALTSANGVILGVEHG
          LNTQITFAVLESFPAFLGLPYDHAMVVAOLGMKLAITMEFVAMLDLKNKLSLPHPT
          VAVATFTLSFANFSTVGMIYGTNSIIDGKSTVIGRWVWMLVSGIAVSLLSAIV
          GLFW"
          6462..6489
          /gene="yxjA"
          complement(6925..7167)
          /gene="yxjA"
          complement(6925..7167)
          /gene="yxjA"
          /function="unknown"
          /codon_start=1
          /transl_table=11
          /protein_id="CAB15939.1"
          /db_xref="GI:2636449"
          /db_xref="SWISS-PROT:P42311"
          /translation="MKNNMLKAGIAVLFSVFAVYAAPSLKAVQAKTPTVSTHYKK
          IKELTYPOVHHVGNAAPEKTIQNELKAIMDRIRNT"
          complement(7290..7306)
          /gene="yxjS"
          complement(7290..7602)
          /gene="yxjS"
          complement(7312..7602)
          /gene="yxjS"
          /function="unknown"
          /codon_start=1
          /transl_table=11
          /protein_id="CAB15940.1"
          /db_xref="GI:2636450"
          /db_xref="SWISS-PROT:P42310"
          /translation="MNRKOLTEAVERDENMMILVPAQCVNHDLDPMELYAKAYQOK
          LNSIKTKTMDIYVPKHAEHIIPQTYIAVLEMGNTDLAQLHSAISGRKO"
          complement(7652..9712)
          /gene="KatB"
          complement(7652..9712)
          /gene="KatB"
          /EC_number="1.11.1.6"
          /note="alternate gene name: Kate"
          /codon_start=1

terminator
gene
CDS
gene
CDS
Query Match.      3.2%; Score 42.4; DB 1; Length 215534;
Best Local Similarity 64.0%; Pred. No. 10;
```

```
Matches 64; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 77 ttaccgaagtcgctgtgcttgccgaagaatatgtcgcatgccccctgataaataca 136
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122719 TTGCCGACATCGGCGTGGATGACGTCGACATTTTCACGACGCTTTTTCCTTGAAAGCTAA 122660
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 137 aggaagtgtaagcgcgtcaaaaaagccaagtgtgttga 176
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122659 AGGAAGGCGATACCGGCAAAAAAGCGAAGTGCCTTTGTA 122620

RESULT 23
AF429315 125020 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
DEFINITION AF429315
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 125020)
            Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
            Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
            Potter,N.T., Ross,C.A. and Margolis,R.L.
            A repeat expansion in the gene encoding junctophilin-3 is
            associated with Huntington disease-like 2
            Nat. Genet. 29 (4), 377-378 (2001)
TITLE 2 (bases 1 to 125020)
JOURNAL MEDLINE 21583737
PUBMED 11694876
REFERENCE 2 (bases 1 to 125020)
            Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.
            Direct Submission
            Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
            Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
AUTHORS
JOURNAL
FEATURES
source 1. 125020
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="16"
            /map="16q24.3; between D16S520 and WI-12410"
            /note="isolated from a patient with Huntington's
            Disease-Like 2 (HDL2)"
            complement(3581..35746)
            /rpt_type=tandem
            /rpt_unit=ctg
            /gene="JPH3"
            complement(<36507..>36887)
            /product="junctophilin 3"
            /gene="JPH3"
            complement(<36507..>36887)
            /gene="JPH3"
            /note="JPH3"
            complement(<36507..36887)
            /gene="JPH3"
            /note="component of the junctional complex between plasma
            membrane and endoplasmic reticulum"
            /codon_start=1
            /product="junctophilin 3"
            /protein_id="AA140941.1"
            /db_xref="GI:17646245"
            /translation="MSSGRNFNDGGSYCGWEDGKAHGHGVCYTGPKGGEGYGSNS
            HGEFVLGYVTWPGSNTYQGTWAGGKRHIGLIESKGVWYKGEWTHGKRGYVRECAQ
            NGAKYEGTSMNGLDGCGTETYSDG"
BASE COUNT 29056 a 32731 c 30696 g 28283 t 4254 others
ORIGIN

Query Match      3.0%; Score 40.2; DB 9; Length 125020;
Best Local Similarity 10.6%; Pred. No. 35;
Matches 82; Conservative 315; Mismatches 367; Indels 6; Gaps 1;
QY 303 gtccgaagcgtacctgaagcgtgtgcgaatatgagcagcaaaagtgcgcgca 362
```

```

Db 17821 RKMWRNRYSMRRCAGMGAACGASRSMKCMGGRYKRSKSGSRGTGYMKKGGKGS 17762
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 363 cctgattcaatcagcgttatgtactgcttcgcgaaccgcttcgttcagcaaatccctgc 422
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 17761 MSSRKMKGSSSTRRGRGSAKSCSYMMGMSMCMSMAKSYMRYCMYRMSSSYMK 17702
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 423 cgtatagcgcagcgcgttcgcacatcttcgtaatcgatggacacacccgctgctgc 482
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 17701 CYSCGMSSJSTYSTCCMKSSWSGCTCKMKTYSGMSSYSTMGYISTSCCKYKKSWS 17642
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 483 cgaccctcagcgttcacatcaagaagcgcgcgaagactcaacgcgcgcgttggatc 542
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 17641 MYCKTKSRYRKRKSYIYWGGRKRAKRYKCGRRRBMRYMWCALWMMSYCCMWSCMY 17582
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 543 gaggcgcgtcgcgaagcgttaaatccatgctgttaagcagcagcgcgcgcgtgcgc 602
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 17581 YYSKSCYTKSSTCYKRGYWMSTKCYSNAGKRSRMYYCMNRSSKSSSMASMSWCM 17522
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 603 tgaaatctcgtccaatatcgaaacacatgaatttgcgcgcgcgcgcgcgcgttcgag 662
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 17521 GWGAGYRKRKSKSAGWAGWRSKGRSTOMKRAKSKTKTSYSTGSRMCKKRYSKYSRG 17462
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 663 tgcgcgcacatcattcattcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 722
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 17461 MKGKKTKCYCMWKYKRYKRTSMCMWYMKSMGYKRYKRCMKKCKCTGYRGMSSKRSYK 17402
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 723 ttatcaaacgtgattgctatcgcgcgttctgtctgaacagcgcgcgcgcgcgcgcgc 782
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 17401 SMRGKRGSSYSTSCMKSCGYSWMKCMKYMKYKRRSMKSSGSKRWRYAGRCYSSS 17342
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 783 cgtgcttcgttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 842
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 17341 MWSTKRKRSKCYSYKRGKRMKMGWGMKRGSKYWSSMKMKKRSWSKCYSTKYSKSG 17282
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 843 tgcgaagtgctcaactaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 902
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 17281 RRSKGMGRSTSKAKSSMRMAGSKCTGSSYNNRRNNRMKCTCNYRRSRAMNG 17222
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 903 cgttcgttatgaaacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 956
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 17221 NAAAGCTTCCCCCANTNNGGGGAAAAAGGCGSASASASACYSKYSKYSKYSKYSKYS 17162
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 957 ctaccacacatcagatttcgcttatcgaagaagcgcgcgcgcgcgcgcgcgcgcgcgcgc 1016
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 17161 SKCRGRSRRKSKMGRGRSGGTSYSAKSGRCYCMWGRKGRKCMSSRKMKMYRSY 17102
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 1017 tgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1066
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 17101 RRRKRWMTCKMKCYSMYMYCRSMCMCMKSCCGCYSMGMSYSYSGKY 17052
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 24
BHTUL/c 37000 bp DNA linear VRL 14-AUG-1996
DEFINITION BHTUL bovine herpesvirus type 1 UL22-35 genes.
ACCESSION Z78205
VERSION 278205.1 GI:1491620
KEYWORDS capsid protein; DNA binding protein; DNA polymerase; envelope protein; glycoprotein B; glycoprotein H; proteinase; scaffold protein; tegument protein; thymidine kinase; UL22; UL23; UL24; UL25; UL26; UL26.5; UL27; UL28; UL29; UL30; UL31; UL32; UL33; UL34; UL35; UL36.
SOURCE Bovine herpesvirus 1.
ORGANISM Bovine herpesvirus 1.
VIRUSES: dsDNA viruses, no RNA stage; Herpesviridae;
Alphan herpesvirinae; Varicelloviruses.
REFERENCE 1 (bases 1 to 37000)
AUTHORS Schwyzer,M., Vlcek,C., Lowery,D.E., Bello,L.J., Meyer,G., Mista,V., Thiry,E., and Paces,V.
TITLE Gene contents in a 37-kb segment centered in the UL part of the bovine herpesvirus 1 genome
JOURNAL Unpublished

```

```

REFERENCE
AUTHORS 2 (bases 24211 to 27636)
TITLE Whitbeck,J.C., Bello,L.J. and Lawrence,W.C.
JOURNAL Comparison of the bovine herpesvirus 1 gI gene and the herpes simplex virus type 1 gB gene
MEDLINE Journal of virology. 62 (9), 3319-3327 (1988)
88300884
PUBMED 2841484
REFERENCE
AUTHORS 3 (bases 24215 to 27793)
TITLE Mista,V., Nelson,R. and Smith,M.
JOURNAL Sequence of the bovine herpesvirus type-1 glycoprotein gene that is homologous to the herpes simplex gene for the glycoprotein gB
MEDLINE Virology. 166 (2), 542-549 (1988)
89020821
PUBMED 2845660
REFERENCE
AUTHORS 4 (bases 27788 to 29695)
TITLE Haanes,E.J., Thomsen,D.R., Martin,S., Homa,F.L. and Lowery,D.E.
JOURNAL The bovine herpesvirus 1 maturational proteinase and scaffold proteins can substitute for the homologous herpes simplex virus type 1 proteins in the formation of hybrid type B capsids
MEDLINE Journal of virology. 69 (11), 7375-7379 (1995)
96013856
PUBMED 7474173
REFERENCE
AUTHORS 5 (bases 31487 to 32728)
TITLE Whitbeck,J.C., Lawrence,W.C. and Bello,L.J.
JOURNAL Characterization of the bovine herpesvirus 1 homolog of the herpes simplex virus 1 UL24 open reading frame
MEDLINE Virology. 200 (1), 263-270 (1994)
94174724
PUBMED 8128626
REFERENCE
AUTHORS 6 (bases 32460 to 33539)
TITLE Bello,L.J., Whitbeck,J.C. and Lawrence,W.C.
JOURNAL Sequence and transcript analysis of the bovine herpesvirus 1 thymidine kinase locus
MEDLINE Virology. 189 (2), 407-414 (1992)
92351539
PUBMED 1322582
REFERENCE
AUTHORS 7 (bases 33521 to 36452)
TITLE Meyer,A.L., Petrovskis,E.A., Dufius,W.P., Thomsen,D.R. and Post,L.E.
JOURNAL Cloning and sequence of an infectious bovine rhinotracheitis virus (BHV-1) gene homologous to glycoprotein H of herpes simplex virus
MEDLINE Biochimica et biophysica acta. 1090 (2), 267-269 (1991)
92031707
PUBMED 1657187
REFERENCE
AUTHORS 8 (bases 36100 to 37000)
TITLE Vlcek,C., Benes,V., Lu,Z., Kutish,G.F., Paces,V., Rock,D., Letcworth,G.J. and Schwyzer,M.
JOURNAL Nucleotide sequence analysis of a 30-kb region of the bovine herpesvirus 1 genome which exhibits a colinear gene arrangement with the UL21 to UL4 genes of herpes simplex virus
MEDLINE Virology. 210 (1), 100-108 (1995)
95313343
PUBMED 7793062
REFERENCE
AUTHORS 9 (bases 1 to 644)
TITLE Schwyzer,M., Styger,D., Vogt,B., Lowery,D.E., Simard,C., Laboisiete,S., Mista,V., Vlcek,C. and Paces,V.
JOURNAL Gene contents in a 31-kb segment at the left genome end of bovine herpesvirus-1
MEDLINE Vet. Microbiol. (1996) In press
95313362
PUBMED 7793062
REFERENCE
AUTHORS 10 (bases 1 to 37000)
TITLE Direct Submission
JOURNAL Submitted (13-AUG-1996) Martin Schwyzer, Institute of Virology, Faculty of Veterinary, Medicine, University of Zurich, Winterthurerstr. 266a, Zurich, CH-8057, Switzerland
MEDLINE Location/Qualifiers
PUBMED 1. 37000
FEATURES
SOURCE /organism="Bovine herpesvirus 1"
/db_xref="taxon:10320"
/clone="HindIII fragments E (partial), G, and A (partial)"
misc_feature 2..7
/note="PstI site located at base 30802. .30807 from left

```


TITLE	Promoter, spliced leader, and coding sequence for BICP4, the largest of the immediate-early proteins of bovine herpesvirus 1	source	1. .135301 /organism="Bovine herpesvirus type 1.2" /db_xref="taxon:79890"
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	3 (bases 111028 to 114234; 123984 to 127190) Schwyzer,M., Wirth,U.V., Vogt,B. and Fraefel,C. BICP2 of bovine herpesvirus 1 is encoded by a spliced 1.7 kb RNA which exhibits immediate early and late transcription kinetics J. Gen. Virol. 75 (Pt 7), 1703-1711 (1994)	repeat_region	235. .430 /rpt_type=DIRECT /rpt_unit=235. .248
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	4 (bases 66900 to 96900) Vlcek,C., Benes,V., Lu,Z., Kutish,G.F., Paces,V., Rock,D., Letchworth,G.J. and Schwyzer,M. Nucleotide sequence analysis of a 30-kb region of the bovine herpesvirus 1 genome which exhibits a colinear gene arrangement with the UL21 to UL4 genes of Herpes simplex virus Virology 210 (1), 100-108 (1995)	mRNA	466. .1280 /note="transcribed from alternative late promoter" /product="RL1 (Circ)" /evidence=experimental
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	5 (bases 1 to 31444) Schwyzer,M., Styger,D., Vogt,B., Lowery,D.E., Simard,C., Laboisiere,S., Mista,V., Vlcek,C. and Paces,V. Gene contents in a 31-kb segment at the left genome end of bovine herpesvirus-1 Vet. Microbiol. 53 (1-2), 67-77 (1996)	CDS	486. .1229 /note="immediate-early and late myristylated protein (villon)" /codon_start=1 /product="circ protein" /protein_id="CAA06078.1" /db_xref="GI:2653292"
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	6 (bases 30801 to 67800) Schwyzer,M., Vlcek,C., Lowery,D.E., Bello,L.J., Meyer,G. and Mista,V. Gene contents in a 37-kb segment centered in the UL part of the bovine herpesvirus 1 genome: the last gap Unpublished	polyA_signal	/db_xref="SPRMBL:Q01342" /translation="MGARASAPAGPPPAHVALDLALSGCTIDLRGGDAVYVSCPT RPVYHMRGRGTAHTPVHFGRAVALICRFEMLYLMKGAVYGEPTGLHRLDS LHDFLTAGLOORDLCLDVLVDQMPVPTFTPEILTELADAPAFPPPSARARS TLRRASMRPAPRTCPHOLAGESIIDSICSPQAAAPGCSLIPACDSGADAGCDAGE TAPDCTAARAPSPGALSRYSSVSVF"
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	7 (bases 96901 to 99695) Letchworth,G.J. and Kutish,G.F. DNA sequence of the BHV-1 UL1 to UL3.5 genes Unpublished	gene	1259. .1264 /gene="Circ" 1259. .1264 /gene="Circ"
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	8 (bases 114235 to 122983) Goltz,M., Buhk,H.J., Broil,H., Lewin,M., Mankertz,A., Boerner,B., Borchers,K. and Weigelt,W. Nucleotide sequence of the HindIII O and K fragments located in the US region of the bovine herpesvirus 1 genome Unpublished	gene	complement(1649. .1654) /note="UL53; UL54" complement(1649. .1654) /gene="UL54"
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	9 (bases 121402 to 123983) Schwyzer,M. Glycoprotein E and US9 genes of BHV1 Unpublished	polyA_signal	complement(1658. .2860) /gene="UL54" complement(1658. .2860) /gene="UL54"
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	10 (bases 1 to 135301) Lowery,D.E., Simard,C., Bello,L.J., Thiry,E. and Vlcek,C. Complete DNA sequence of bovine herpesvirus 1 Unpublished	gene	/note="early BICP2 affecting mRNA processing (cell nucleus)" /codon_start=1 /product="UL54 protein" /protein_id="CAA06079.1" /db_xref="GI:2653293"
JOURNAL MEDLINE REFERENCE AUTHORS TITLE	11 (bases 1 to 135301) Schwyzer,M. Direct Submission Submitted (08-SEP-1997) Institute of Virology, Faculty of Veterinary Medicine, University of Zurich, Winterthurerstrasse 266a, Zurich CH-8057, Switzerland Location/Qualifiers 1. .135301	CDS	/db_xref="SPRMBL:Q01354" /translation="MADPEIATLSTSESDDLSFGSDREDEDEAPSLAPALSVQO VKKRLLEGAEDEPMPEPPGEGASGDGPAPRAARVRRRRRRRRRQAPAE ORSRGPAARREALATSSHGGAARSIGSLRLASIAEAORAEVTVAFGA RDLMPVONGEFGRAAGVSPMAVIDFGAEQFVPEGRVYMETLRFHGPDLVMEFVR PRAAQAARLRDLVLRANLVNLAASADSCITMCKFLATKRLRPDIPVAAGAVL ENLRKLAFLLKCYLGRGLPSLEICARRSLATCPASYSMTMLARLSRAVSGAE CVPLLEVTVGDAPFEERYIPGTCVAGLDLADLTHKQACDSMTCLVANFTLVPMHCK YYCNDIEF"
FEATURES source	Complete DNA sequence of bovine herpesvirus 1 Unpublished	gene	complement(3040. .4038) /gene="UL53"
FEATURES source	Complete DNA sequence of bovine herpesvirus 1 Unpublished	CDS	complement(3040. .4038) /gene="UL53"
FEATURES source	Complete DNA sequence of bovine herpesvirus 1 Unpublished	gene	/codon_start=1 /product="glycoprotein K"
FEATURES source	Complete DNA sequence of bovine herpesvirus 1 Unpublished	CDS	/protein_id="CAA06080.1" /db_xref="GI:2653294"
FEATURES source	Complete DNA sequence of bovine herpesvirus 1 Unpublished	gene	/db_xref="SPRMBL:O65816"
FEATURES source	Complete DNA sequence of bovine herpesvirus 1 Unpublished	gene	/translation="MLGGRTVALALALTATLALALWVALAARORCAVPRATANG GSLRMELRSPGAYVWGANNATLADAPCRHAAVVOHITPGLLDGDALGRYRAVAG ARDCAVYLMCAOARGGLAMLVAVAVYLMQERRMGLCHNDLDFSPGTYVAAA ALAAVVGHPYTRKLAMELSARRRLAVDFLDPDLCAMRRAALPLLAEGPARAG ARIAAGSGVITRHPCAAAAYPLVLIKIMAWVAFAGLEISLRYRRRRGGCAGDG GDBGESGIRKVCYNCSTLLAGLVVALVLAIVGVIALHHEHNRLLNLAAQT"
FEATURES source	Complete DNA sequence of bovine herpesvirus 1 Unpublished	gene	complement(4013. .7237) /gene="UL52"
FEATURES source	Complete DNA sequence of bovine herpesvirus 1 Unpublished	CDS	complement(4013. .7237) /gene="UL52"
FEATURES source	Complete DNA sequence of bovine herpesvirus 1 Unpublished	gene	/codon_start=1 /product="component of DNA helicase/primase complex"
FEATURES source	Complete DNA sequence of bovine herpesvirus 1 Unpublished	CDS	/protein_id="CAA06081.1"

TITLE	Potter,N.T., Ross,C.A. and Margolis,R.L.
JOURNAL	A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2
MEDLINE	Nat. Genet. 29 (4), 377-378 (2001)
PUBMED	21583737
REFERENCE	11694876
AUTHORS	2 (bases 1 to 125020)
TITLE	Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.
JOURNAL	Direct Submission
FEATURES	Submitted (05-OCR-2001) Psychiatry, Johns Hopkins Medical Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
source	Location/Qualifiers 1..125020

BASE COUNT	29056	a	32731	c	30696	g	28203	t	4254	others
repeat_region	/chromosome="16" /map="16q24.3; between D16S520 and WI-12410" /note="isolated from a patient with Huntington's disease-like 2 (HDL2)" complement(35581..35746) /rpt_type=candem /rpt_unit=Ctg complement(<36507..>36887) /gene="JPH3" /product="junctophilin 3" complement(<36507..>36887) /gene="JPH3" /note="Jp3" complement(<36507..36887) /gene="JPH3" /note="component of the junctional complex between plasma membrane and endoplasmic reticulum" /codon_start=1 /product="junctophilin 3" /protein_id="AAL40941.1" /db_xref="GI:17646245" /translation="MSSGGRRNPDDGSGSYGGGWDGKAHGHGCTGKPGKGCEYIGSSMS HGPEVLGYTPMSGNTYQGTWAGKRGIGLESgKwVYGEWTHGFGKRGYRGREACAG NGAYESTWNSGLDDGTYGTEYSNG"									
gene										
CDS										

[illegible][illegible]

RESULT	28
AB047272	
LOCUS	9918 bp DNA linear BCT 14-FEB-2001
DEFINITION	Pseudomonas sp. CA10 cat operon (catR, catB, catC, catA, ORFcat5,
ACCESSION	ORFcat6, ORFcat7, ORFcat8), complete cds.
VERSION	AB047272.1
KEYWORDS	AB047272.1 GI:12862416
SOURCE	
ORGANISM	Pseudomonas sp. CA10 (strain:CA10) DNA.
	Pseudomonas sp. CA10
	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
	Pseudomonas.
REFERENCE	1 (sites)
AUTHORS	Nojiri,H., Maeda,K., Sekiguchi,H., Nakai,S., Yoshida,T., Hiroshi,H.
TITLE	and Omori,T.
JOURNAL	Analysis of carbazole degradation genes from pseudomonas sp. strain
REFERENCE	CA10
AUTHORS	Unpublished
TITLE	2 (bases 1 to 8918)
JOURNAL	Nojiri,H., Omori,T. and Habe,H.
REFERENCE	Direct Submission
AUTHORS	Submitted (10-AUG-2000) Hideaki Nojiri, The University of Tokyo,
TITLE	Biotechnology Research Center; 1-1-1 Yayoi, Bunkyo-ku, Tokyo
JOURNAL	113-8657, Japan (E-mail:anojiri@mail.ecc.u-tokyo.ac.jp,
FEATURES	Tel:+81-3-5841-3064, Fax:+81-3-5841-8030)
SOURCE	Location/DnaIifiers 1..8918

gene
CDS

gene
CDS

gene
CDS

```
/organism="Pseudomonas sp. CA10"  
/strain="CA10"  
/db_xref="taxon:135214"  
complement(594..1475)  
/gene="catR"  
complement(594..1475)  
/gene="catR"  
/function="Positive regulator of cat operon"  
/codon_start=1  
/transl_table=11  
/product="positive regulator of cat operon Catr"  
/protein_id="BAB32455.1"  
/db_xref="GI:12862417"  
/translation="MEDRLRYFOALAEFLNFTFAERLHIAOPLSRIOIOLDELDG  
VLAELGRPLRTTEAGRFHFHSSTILETLKVDNRIRSGEKATIGTGFASPTTY  
GVLPDLIRLRSPAGLELGSEMTTILOQVESLKAGRIDVCGRHINDPAIQOOVTYI  
DPYVALPRAGHPLLPGFVSITLDLAQESFVLIPGNRPISADHVIALTPAAGLQIEVAQ  
WNLDTAICGLVAAGIGITLVPAVOYLHRADIGYAKLQEOATSPIITSRTDASE  
ATWCHCALIBELSRRGL"
```

1612..2733
/gene="catB"
1612..2733
/gene="catB"
/function="cis,cis-muconate cycloisomerase"
/codon_start=1
/transl_table=11
/product="cis,cis-muconate cycloisomerase Catb"
/protein_id="BAB32456.1"
/db_xref="GI:12862418"
/transl_table="MSOVLETSEVEETIIIVDLLPTRPHKLAMHTMQHOTLVYIRLCSOG
ICGLGSTTGIGLAYNESESPEIKONIDSHLAPLLIGODAAANTIINAMLRIIDKARKGMT

ARTADVAATOVLSVLEDERDADDCPELTGIALTVLALNQAALGAGGRUATRG
AVSTNPADPVTHPVQAAMGLGRVLEHPRIMGLVDLPQFEDRACORLGIITAVK
DAPGEDOVALRATGVSGRRLVRLTVEALPTPAETATVGVITGCTGLGEVAVRWL
ARAGAOHLVLTSRGPDPAAGAELEAGGSPSVACVADADALADREDFVLGSI
LPLTVGVTAGVGVGPIDITLSTAEAGTAAKLAGLADLADREDFVLGSI
ACWAGSGNOSVGAANAYIDALALHRRARGLAITSVANGPMAEAGMAADAVSETLR
OGGLDLPAPAMTELRAVYVODTVYADVMORVAPLTSRPSALIGLPEVAL
AADERTQDADGASEVTVTRALAEPEOBLRLTDLVRESATVGLHSSADVPEGAP
RDVGEDSLTAVELKRLGAAATGLSPSTMVEDYPPLELAQLRLHLLAVLEAGVAP
ATGGADEPIAIIIGMACFPGVSPBOLMDLAVAGTGAISEPVNRCWOTHLDEPD
PDPRGTYSTOGGHEADEFPPTFGISPREALVMDPOQLLETWESFEFAGIRP
ETLSTGTGSSGYOEGGAGGEGTHVAGSPSLRSLVFLGEPVAVVD
TACSSIVAHILACOSLPGNSNLAVAGATITMTNPNPLASRORALAKDRCKAFS
DDADGNTLAEVGVYVRLSDAQRNGHVLAVLNGSAINODGASPEGLTAPNGSQR
VIRIOLANARLAPDIDLEAHGTGTPLDLEAOLAFVTRGRDRRBEALILGSKS
NIGHTOSAAIASVAKMYMALHSELPLPLHADAPSHEVRAVAVVPPAGTPIALR
GRPRRAVSSFGISGTNAHVLEQAPVADTPAEERPAVPAVIAAGVPMVYARSA
ALRGOAERLHAETVGTALPAAGPLDGLSLVSNAREHREAVVPPAGTPIALR
AVADGSPVAVRGVADVEGRVYEPVPGSGWQVSGSLDDESAFARIECAAL
AEFTDMSLVVLRGVGAPSLERVDVNPASFAVWVSLAIIHRSRGLVDVAVVSGG
ETIAAVVSGALSLRGRVAVLRSGAIGALRGMSVALLSVYLERIVFEBCRV
SVAAVNGRVSVAAGEPALDALARLADTARLADVAVSHSHQVEDLHELLEV
LAELAPRTSEVPFSTVGDMDLTARMAGVFRMLRGVRRADVAADLAAETRAV
EVSSHVPLSMVVOEAIDEAGVPAVAGTLRDOGGTDFRLLSAAEFVGVVDVWGL
FEGTGAARLIDLPYAFHQBHLMAVPAVPAVAAADDDAFAFVAVEDGVSAITLALG
TDEBSVAVVPLATISMRARRDRSTVDAMRVAVMPLGCTLPHPELTGTMLVTADG
IDDDVAGALETIGAEVRRLVLEDECVDRVAVRERLAGADVTGIVSVLAAARTAV
PGTSLVIGTALITVALIQALGDAIDAPVVALRGAVSTGRADLTLPVQAOVGTIGT
ALBHPORWGTLDLPALADARAQRLAVALGSLGALSDQQLARPGVFRIRIVAEV
TAGRPAGTWTPTGTTLVGSGSLTAPHLARWLARQAEHLVLISRGTAAPGAELIVA

Query Match 2.9%; Score 38.8; DB 1; Length 125580;
Best Local Similarity 48.6%; Pred. No. 79; Mismatches 112; Indels 0; Gaps 0;
Matches 106; Conservative 0;

Qy 1104 caacgagcgagccgcgcacatggtacgcagcactatgacgagcgttaatgcccgttga 1163
Db 49632 CACCGCGCGGCTCTCGACGACGCCACCGTGGACCGCTGACCGCGGCTGACAC 49631
Qy 1164 catctgctactactcttcttgcgcgatttaactgctgcgcgacacgagcgagc 1223
Db 49692 CCTCTGCGCGCCGACGACGCCGCTTGGACCTGACGACGCCACCGCGACCGCGA 49751
Qy 1224 ttgaggtcttggaattggaagaagaccctgcttgcagcttcgtcccgag 1283
Db 49752 CCTGCGAGGCTTGTGCTGTACCTCTCGGTCGCCGCTGACCGCGCGCGCCGACAGG 49811
Qy 1284 caaatagcaatagcgccgcgttgcgcaagtgcgtg 1321
Db 49812 CAACATACGCGCGCGACACGCTTCTCGACGCGCTCG 49849

RESULT 30

AX211739 125401 bp DNA linear PAT 06-SEP-2001
LOCUS Sequence 35 from Patent WO0159126.
DEFINITION AX211739
ACCESSION AX211739
VERSION
KEYWORDS GI:15523950
SOURCE
ORGANISM
Streptomyces noursei.
Streptomyces noursei.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
1 (bases 1 to 125401)
Zotchev,S.B., Sekurova,O.N., Fjervik,E., Brautaset,T.,
Stroem,A.R., Valila,S., Ellingsen,T.E., Sletta,H.V. and
Gulliksen,O.M.
Gene cluster encoding a nystatin polyketide synthase and its
manipulation and utility
Patent: WO 0159126-A 35 16-AUG-2001;
Norges Teknisk Naturvitenskapelige Universitet (NO) : STIFTETSEN
IND OG TEKNISK FORSKNING VED NORGES TEKNISKE HOGSKOLE (NO) :
ALPHARMA AS (NO) : SINVENT AS (NO) : Zotchev, Sergey Borisovich

FEATURES
source location/Qualifiers
1. 125401
/organism="Streptomyces noursei"
/db_xref="taxon:1971"
/note="ATCC 11455"
BASE COUNT 15664 a 49692 c 42871 g 17174 t
ORIGIN

Query Match 2.9%; Score 38.8; DB 6; Length 125401;
Best Local Similarity 48.6%; Pred. No. 79; Mismatches 112; Indels 0; Gaps 0;
Matches 106; Conservative 0;

Qy 1104 caacgagcgagccgcgcacatggtacgcagcactatgacgagcgttaatgcccgttga 1163
Db 50040 CACCGCGCGGCTCTCGACGACGCCACCGTGGACCGCTGACCGCGCTGACAC 50099
Qy 1164 catctgctactactcttcttgcgcgatttaactgctgcgcgacacgagcgagc 1223
Db 50100 CTTCTGCGCGCCGACGACGCCGCGCTGGACCTGACGACGCCACCGCGACCGCGA 50159
Qy 1224 ttgaggtcttggaattggaagaagaccctgcttgcagcttcgtcccgag 1283
Db 50160 CTTCTGCGCGCCGACGACGCCGCGCTGGACCTGACGACGCCACCGCGACCGCGA 50219
Qy 1284 caaatagcaatagcgccgcgttgcgcaagtgcgtg 1321
Db 50220 CAACATACGCGCGCGACACGCTTCTCGACGCGCTCG 50257

RESULT 31

AE005062 10500 bp DNA linear BCT 12-FEB-2001
LOCUS Halobacterium sp. NRC-1 section 93 of 170 of the complete genome.
DEFINITION AE005062 AE004437
ACCESSION AE005062.1 GI:10580956
VERSION
KEYWORDS
SOURCE
ORGANISM
Halobacterium sp. NRC-1.
Halobacterium sp. NRC-1.
Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
Halobacteriaceae; Halobacterium.
REFERENCE
AUTHORS
1 (bases 1 to 10500)
Ng,M.V., Kennedy,S.P., Mahairas,G.G., Bergquist,B., Pan,M.,
Shukla,H.D., Lasky,S.R., Balliga,N., Thorsson,V., Sbrogna,J.,
Swartzell,S., Weir,D., Hall,J., Dahl,T.A., Weili,R., Goo,Y.A.,
Leithausen,B., Keller,K., Cruz,R., Danson,M.J., Hough,D.W.,
Maddocks,D.G., Jablonski,P.E., Krebs,M.P., Angevine,C.M., Dale,H.,
Isenbarger,T.A., Peck,R.F., Pohnschrod,M., Spudis,J.L.,
Jung,K.-H., Alam,M., Freitas,T., Hou,S., Daniels,C.J., Dennis,P.P.,
Omer,A.D., Ehardt,H., Lowe,T.M., Liang,P., Riley,M., Hood,L. and
Dassarma,S.
From the cover: genome sequence of halobacterium species NRC-1
Proc. Natl. Acad. Sci. USA 97 (22), 12176-12181 (2000)
11016950
2 (bases 1 to 10500)
Ng,M.V., Kennedy,S.P., Mahairas,G.G., Bergquist,B., Pan,M.,
Shukla,H.D., Lasky,S.P., Balliga,N., Thorsson,V., Sbrogna,J.,
Swartzell,S., Weir,D., Hall,J., Dahl,T.A., Weili,R., Goo,Y.A.,
Leithausen,B., Keller,K., Cruz,R., Danson,M.J., Hough,D.W.,
Maddocks,D.G., Jablonski,P.E., Krebs,M.P., Angevine,C.M., Dale,H.,
Isenbarger,T.A., Peck,R.F., Pohnschrod,M., Spudis,J.L.,
Jung,K.-H., Alam,M., Freitas,T., Hou,S., Daniels,C.J., Dennis,P.P.,
Omer,A.D., Ehardt,H., Lowe,T.M., Liang,P., Riley,M., Hood,L. and
Dassarma,S.
Direct Submission
Submitted (14-JUL-2000) Institute for Systems Biology, 4225
Roosevelt Way NE, Seattle, WA 98105, USA
location/Qualifiers
1. 10500
/organism="Halobacterium sp. NRC-1"

```
gene /strain="NRC-1"
CDS /db_xref="taxon:64091"
      /complement(86..232)
      /gene="VNG1459h"
      /complement(86..232)
      /note="hypothetical protein"
      /codon_start=1
      /transl_table=11
      /product="Vng1459h"
      /protein_id="AAG19765.1"
      /db_xref="GI:10580957"
      /translation="MDLSVRADDKGMGVCECHGPVEKIDPWEGATRIEGRKGEV
      YHEV"
gene complement(522..734)
CDS /gene="VNG1461h"
      /complement(522..734)
      /gene="VNG1461h"
      /note="hypothetical protein"
      /codon_start=1
      /transl_table=11
      /product="Vng1461h"
      /protein_id="AAG19766.1"
      /db_xref="GI:10580958"
      /translation="MWMHIVEDAGSGSGMEKQDLVAVHVAVALIAFGVLRISRG
      QVRGALNGLPAIVVGVGVKRYOLV"
gene complement(734..2809)
CDS /gene="cdca4a"
      /complement(734..2809)
      /gene="cdca4a"
      /note="cdca4a"
      /codon_start=1
      /transl_table=11
      /product="cell division cycle protein"
      /protein_id="AAG19767.1"
      /db_xref="GI:10580959"
      /translation="MTRVRYTLRDADRTADAPAVPGQLSFTAETRQOIGVAEGDAV
      TYGSEDPVAVPGPTTAAASSGLAGAVLNRVGPDTVAVAPADPSPAELTVAPA
      TSTTSGADALRALAGRHVLAADDRIIVLPSGVVTVLDVEDTGGPGVMTAPTAA
      ITLTGHAOSGHADAHPATRTGCGIDDESGALIRLVAVPVDASVYATICVTPAG
      VLVHSPAGTGTTLVRAVAAAADLVESVAPEDAGDRDLAAVLDAARAPBCQVYFV
      ESLAAPDPTADGASGRSPSALGMLDVRGHDVTVVVGETDDADVPALRRGGR
      FDAEVRVGPDPAAARRAIIIDVHTDGVRLADAVSLAVADRTGCTGADLTAVLDVAT
      RAAGSAGSPVIRORDLEALDVAAGPSTLRDASVQPTTYODIGLDRAKREVVTV
      EMPQRYPALFERLDAAAPTVGLHGGPGTKMLAKAVASTDANFLSYDGPELMARY
      VGESERGVLDLFEARARLAPVAVFLDEVDLSIPARHDTDTGASERVVSQLTLELDLS
      PRGSVAVLATNRRESVDPALRPRIETQVAVPIPDODARAAIFEVOLDGVAATGDS
      TTAALAAATGYTGSIDAGVYRSGALLAMEDHLRETFEDATDSGLVTVQRHLRAIET
      TAPSYDAGC"
gene 3000..3428
CDS /gene="bip"
      /note="VNG1463G"
      /complement(3000..3428)
      /gene="bip"
      /note="bip"
      /codon_start=1
      /transl_table=11
      /product="bacterio-opsin linked product"
      /protein_id="AAG19768.1"
      /db_xref="GI:10580960"
      /translation="MTVEQOATIDAPPGDVFSEFLDPPRNHVKITPGLITVVDVEALA
      SGGRAKRYKLAGVELVGTVDIEHREPRRLVQELTGAIDGRTYRSLGDDTSTSLTY
      EAEYDLPDVIVSVLARIKAYNKRKAATVNTLTFLFH"
gene complement(3462..5486)
CDS /gene="bat"
      /note="VNG1464G"
      /complement(3462..5486)
      /gene="bat"
      /note="bat"
      /codon_start=1
      /transl_table=11
      /product="bacterio-opsin activator"

gene /protein_id="AAG19769.1"
CDS /db_xref="GI:10580961"
      /translation="MTSYONTSESETAAGATTIGVLEFAGSDPFGTAPACDLDEEDGRFDV
      TQIRPVTAVARDVDDPDIDCYAVVAREDDPECPVAFLEAVRQTHAEPVAVVPAVED
      VARRAVDADATGIVPAVSEDAITRAITDRIEQSNAPASEDETERMPISDLTVSEERLK
      EQLADLPAGITITSDATDEEPTIYINDSEFIDTIGSPUEVVGANHRFLOGPRTNDR
      VAEFWTAIREDHDTOYVLRNRYRDSGLFVNQVDISPIYDEDTVSHYVGFOMDVSER
      AAOEELQGEROSIDRLDRVNGIMNDVTSALVAGAVAREIETRIIDRIOTGGEYAGAV
      PCRVDATEDTITVAEAGDCEGDCDGPVPLASAEVALLQDVEQREALVSTDADGV
      SGTADGDACVAVLVYRSTTYGLAVSTAENRIDDEOYVLRSLGRTGASINDALTR
      RITADTVNLNIGVELSDTALFLVELAGATDTFEQPAITADSGOTGLMLVYTPHPDDP
      QAVVTALCYDAQDAVEYVSTDDESVPVFDLSSSPLNYSRCSGRVLRMARADPTL
      ELDRVGTGGAARRVLTDRKTDADVELVATIHEDDEBOTPHGRREBLNDLDRQLTA
      LQKAVYSGFEEWPRAREGKQLAESMDIVPSTYHQHQAOKLVGAFFEE"
gene complement(5483..6589)
CDS /gene="bip"
      /note="VNG1465G"
      /complement(5483..6589)
      /gene="bip"
      /note="bip"
      /codon_start=1
      /transl_table=11
      /product="bacteriorhodopsin related protein"
      /protein_id="AAG19770.1"
      /db_xref="GI:10580962"
      /translation="MGSDVCJHMSRQFVPSWVLEPEAAGDLPLTVSRSLTALAAA
      FAVGAGAGFVAPLEVOAGVYLLGMVAMNPHGXEHEEMLRRRAASFOCKYIYAYLVG
      IAPGALFEFVAVAGLAVTVAVAKGCGGCVOSMDALYCTDHLTRPORMLAAVYRG
      GAVMVVPLMFWDVFEYAFSSWISTDPBASVSLGSDITRRLVLDGGVRSVVAHLG
      LGYRRAAGTGSFLADAEFTLLIAYPALVPAVAVGLDYPPELPMVSAQVARSVAVDTA
      MTOADATGMIDLADDDPARATLASVAVIVSVAFFGLAAVYLMLSLPOLPGGGITLV
      GLVAFWSIVFSIITLPHVVVGGWLDRTGIIWVP"
gene 6711..6845
CDS /gene="VNG1466h"
      /complement(6711..6845)
      /gene="VNG1466h"
      /note="hypothetical protein"
      /codon_start=1
      /transl_table=11
      /product="Vng1466h"
      /protein_id="AAG19771.1"
      /db_xref="GI:10580963"
      /translation="MGLPMGAIYKSTVTAASROEPTSDTOKRVTECRANDHEFVREWM
      "
gene 7089..7877
CDS /gene="bop"
      /note="VNG1467G"
      /complement(7089..7877)
      /gene="bop"
      /note="bop"
      /codon_start=1
      /transl_table=11
      /product="bacteriorhodopsin"
      /protein_id="AAG19772.1"
      /db_xref="GI:10580964"
      /translation="MLELPTAVEGYSQAQITGRPEMIMLALSTALMGITLTVLVKG
      MGVSDPDARKKFAVITIVPAIATMYLIGLGMVPPGSGEONIYVARVADMLFT
      TELLIDIALIWDADGCTIATVAGDGMICGVALTKVYSRYVMAVIASTAAIMY
      IIVLVEGFTSKAESRPEVASTFKVLRNVTVYMGATPYVMHIGSEAGACIVPLNIET
      LLEWLVDSAKGEGFLLIKSRAIFSEAEAPSPSAGCAATSD"
gene 7929..8198
CDS /gene="VNG1468h"
      /complement(7929..8198)
      /gene="VNG1468h"
      /note="hypothetical protein"
      /codon_start=1
      /transl_table=11
      /product="Vng1468h"
      /protein_id="AAG19773.1"
      /db_xref="GI:10580965"
      /translation="MSBPLGIVLARSRDGDADDPASINRLAVVALAGALVGAAG
      LPAVPLRSFGMTFMFAFYVGVSEFVSATVVALAGLHLYTTTDD"
gene 8329..9603
```


Db 7603 GCGCGGATGCC 7593

RESULT 34

SC10A5

DEFINITION Streptomyces coelicolor cosmid 10A5. 38734 bp DNA 1 linear BCT 26-APR-2000

ACCESSION

AL021529.1 GI:2815309

VERSION

ABC-transporter; cold-shock domain protein; DEAD-box RNA helicase;

KEYWORDS

ABC-transporter; cold-shock domain protein; DEAD-box RNA helicase; NTP pyrophosphohydrolase; oxidase; oxidoreductase; periplasmic oligopeptide binding protein; phosphoenolpyruvate-utilizing enzyme; secreted protease; transcriptional regulator; transposase

SOURCE

ORGANISM

1 (bases 1 to 38734) Streptomyces coelicolor A3(2).
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.

REFERENCE

AUTHORS

Redenbach, M., Kleiser, H.M., Denapalte, D., Eichner, A., Cullum, J.,

TITLE

Kinashi, H., and Hopwood, D.A.
A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome

JOURNAL

Mol. Microbiol. 21 (1), 77-96 (1996)

MEDLINE

97000351

REFERENCE

2 (bases 1 to 38734)

AUTHORS

Murphy, L. and Harris, D.

TITLE

Unpublished

JOURNAL

3 (bases 1 to 38734)

REFERENCE

Parkhill, J., Barrell, B.G. and Rajandream, M.A.

AUTHORS

Submitted (26-JAN-1998) Streptomyces coelicolor sequencing project,

TITLE

Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge

JOURNAL

CB10 ISA E-mail: barrell@sanger.ac.uk Cosmids supplied by Prof.

REFERENCE

David A. Hopwood, [3] John Innes Centre, Norwich Research Park,

AUTHORS

Colney, Norwich, Norfolk NR4 7UH, UK

TITLE

Notes:

COMMENT

Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBSRC.

REFERENCE

Details of S. coelicolor sequencing at the Sanger Centre are available on the World Wide Web.

AUTHORS

(URL: <http://www.sanger.ac.uk/Projects/S.coelicolor/>)

TITLE

CDS are numbered using the following system eg SC7B7.01c. SC (S. coelicolor), 7B7 (cosmid name), .01 (first CDS), c (complementary strand).

REFERENCE

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous.

AUTHORS

Usually the length in codons is given for each CDS.

TITLE

The length in codons is given for each CDS.

REFERENCE

Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database.

AUTHORS

The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon.

TITLE

Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al.,

REFERENCE

Nucleic Acids Research, 22(22):4768-4778(1994) and the Frameplot program of Bibb et al., Gene 30:157-66(1984) as implemented at

AUTHORS

<http://www.nih.gov/jp/jun/cgi-bin/frameplot.pl> CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation

FEATURES

source

1. .38734

/organism="Streptomyces coelicolor A3(2)"

/strain="A3(2)"

gene

misc-feature

CDS

/db_xref="taxon:100226"
/clone="cosmid 10A5"
1. .289
/gene="SC10A5.01"
1. .108
/gene="SC10A5.01"
/note="overlap with st37 from: 23623 to: 23730"
1. .289
/gene="SC10A5.01"
/note="SC10A5.01, probable phosphoenolpyruvate-utilizing enzyme, partial CDS, len: 295 aa; overlaps and extends ORF from neighbouring cosmid SC37.16; similar to many eg. PPSA5 TAMA P46893 probable phosphoenolpyruvate synthase (834 aa), fasta scores; opt: 215 z-score: 351.4 E(): 2.3e-12, 47.8% identity in 69 aa overlap"

RBS

gene

CDS

/note="possible RBS upstream of SC10A5.02"
367. .1554
/gene="SC10A5.02"
367. .1554
/note="SC10A5.02, probable oxidase, len: 395; similar in N-terminus to many diverse oxidases eg. TR:P95483 (EMBL:U7449 3) AMINOPYRROLININ OXIDASE PRND from PSEUDOMONAS FLUORESCENS (363 aa), fasta scores; opt: 349 z-score: 372.1 E(): 1.6e-13, 33.3% identity in 165 aa overlap"

gene

CDS

/codon_start=1
/transl_table=11
/label="SC10A5.02"
/protein_id="CA16434.1"
/db_xref="GI:2815311"
/db_xref="SPTRMBL:054095"
/translation="MIPNWPYIVENQVGNKPLGVRRGDIYVLRDIDNLYCGV ARCPKGNLGDGRKRGNTIECFHGFRTGADGACVTRPKSEARIPSSIVPYPVG REQFLVMMWMCDEPRTADLPVAAPEVTDNRKLYATKRWTRPVHTRIESTLEFY HTVYLRDHWFNYIDYLLYLPSPKFGIDGRERYLAATRIITHNRYETAEAGQIRSF DHCOEDDPNTHTHYITFEPCMVHGTQEPFTTSLVPIIDQNTSHILRWYEYEDVK PVIPEPRLIRLLIPASLYMEKVVQDQDVRIHEHOPKISAGVNFIPVDEMNAYI SMRAKLIDASAPSPPARAAPEPBAAGRGSAARATNGRGAAGRGTRKEDAA AKP"

RBS

gene

CDS

/gene="SC10A5.03"
1597. .2277
/gene="SC10A5.03"
1597. .2277
/note="SC10A5.03, probable membrane protein, len: 226 aa"
/codon_start=1
/transl_table=11
/label="SC10A5.03"
/protein_id="CA16435.1"
/db_xref="GI:2815312"
/db_xref="SPTRMBL:054096"
/translation="MYGVDASVGPKALVTAFNTVAAGVWFLFGADTVADMGFD EDAVTLRVLATLSVLYLIRIATNFPMLOKMKMSSEATGIVIVYHGMVYRG ETDAGVSAFTWLVGLYLGSLYNTGSEYORKMKRPNKRLITGEGFKRSMHN YFGDAVLFSGFALVYGTGPWAFALPIIMVCMFVLPIMDKYLAERTGEAFDEYAST AKFVPVY"
2433. .2436
/note="possible RBS upstream of SC10A5.04"
2442. .3329
/gene="SC10A5.04"
2442. .3329
/gene="SC10A5.04"
/note="SC10A5.04, unknown, len: 295 aa; similar to hypothetical proteins from Mycobacterium tuberculosis

Db 363 TCCACACACCCGCTTTCACGCTGTTGGTGAATGCCGACGACGATCATCG 304
Qy 1127 taccgacgcgactatgagcgtaatgscgttggacatcctgcctactgtcttgc 1186
Db 303 TGCAGGCGCGCTGGGACGAGGCTTCGACATCACCGGCTCTGGGTGGCAAAATTC 244
Qy 1187 gcgattatcgttggcgatccgacagcgagcttgggtgttgcgtgaattgagc 1246
Db 243 ATCGGTCATCTGACCAACCGGCTTGGGGTAGTCTGATGTAATTCACCGCTGGGCG 184
Qy 1247 aagaagacctgcct 1260
Db 183 GTGACGACGACGTT 170

RESULT 36
SME591791/c 340900 bp DNA linear BCT 16-AUG-2001
LOCUS Sinorhizobium meliloti 1021 complete chromosome; segment 10/12.
ACCESSION AL591791 AL591688
VERSION AL591791.1 GI:15075538
KEYWORDS
SOURCE Sinorhizobium meliloti.
ORGANISM Sinorhizobium meliloti.
Bacteria: Proteobacteria; alpha subdivision; Rhizobiaceae group;
Rhizobiaceae, Sinorhizobium.
1 (bases 1 to 340900)
Capela,D., Barloy-Hubler,F., Gouzy,J., Bothe,G., Ampe,F., Batut,J.,
Boisard,P., Becker,A., Boutry,M., Cadieu,E., Dreano,S., Gloux,S.,
Godrie,T., Goffeau,A., Kahn,D., Kiss,E., Lelaire,V., Masuy,D.,
Pohl,T., Portetelle,D., Puehler,A., Purnelle,B., Rampsperger,U.,
Renard,C., Thebault,P., Vandenbol,M., Weidner,S. and Galibert,F.
From the Cover: Analysis of the chromosome sequence of the legume
symbiont Sinorhizobium meliloti strain 1021
Proceedings of the National Academy of Sciences of the United
States of America. 98 (17), 9877-9882 (2001)
11481430
2 (bases 1 to 340900)
Gouzy,J.
Direct Submission
Submitted (26-JUL-2001) Gouzy J., Submitted on behalf of the MELLILO
EU Consortium
MELLILO EU Consortium:
Laboratoire de Biologie Moléculaire des Relations
Plantes-Microorganismes, UMR215-CNRS-INRA, BP27, F-31326 Castanet,
France, Laboratoire de Génétique et Développement UMR6061-CNRS,
Faculté de Médecine, 2 avenue du Pr. Léon Bernard, F-35043 Rennes,
France, GATC GmbH, Fritz-Arnold-str. 23, D-78467 Konstanz, Germany,
Universitaet Bielefeld, Biologie IV (Genetik) Universitaetstr 25,
D-33615 Bielefeld, Germany, Unité de Biochimie physiologique,
Université Catholique de Louvain, Place Croix du Sud 2, Bte 20,
B-1348 Louvain-la-Neuve, Belgium, Unité de Microbiologie, Faculté
des Sciences Agronomiques de Gembloux, Avenue Mearchal Juin 6,
B-5030 Gembloux, Belgium. E-mail:terome.gouzy@oulouise.inra.fr
http://sequence.loulouise.inra.fr/meliloti.html.
Location/Qualifiers
1. 340900
/organism="Sinorhizobium meliloti"
/strain="1021"
/db_xref="taxon:382"
132. .518
/gene="SMC01981"
132. .518
/gene="SMC01981"
/function="small molecule metabolism; energy transfer;
electron transport"
/note="Product confidence : putative
Gene name confidence : hypothetical
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11

gene
CDS
/evidence=not_experimental
/product="PUTATIVE CYTOCHROME C TRANSMEMBRANE PROTEIN"
/protein_id="CAC47094.1"
/db_xref="GI:15075539"
/transl_table="MR1ALALSPALSICPVAGLAEGBAGATYFKKATCTHYID
KDONKVSLOGVIGRTAGTTHADFKYSQAMIDAGKGLVMDATLAEVLRNRAKVKG
TKMVFPLKKEDEIANIAYLKQHK"
656. .1600
/gene="coxM OR SMC01982"
656. .1600
/gene="coxM OR SMC01982"
/EC_number="1.9.3.1"
/function="small molecule metabolism; energy transfer;
electron transport"
/note="Product confidence : probable
Gene name confidence : probable
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="PROBABLE ALTERNATIVE CYTOCHROME C OXIDASE
POLYPEPTIDE II TRANSMEMBRANE PROTEIN"
/protein_id="CAC47095.1"
/db_xref="GI:15075540"
/transl_table="MAVVVILVLLAVGSVLPILLSPMWMTPIASNNYIDNTITTFM
ITGATATATYAFMAVCYLFRRRGNTATAYDENRLSGMATGTFGVAAALAGLF
VNDVETVPDASAEVYIGOWMSFRLPGADKRGITETDIAPENLGVRRDAG
QDDIILIGSELHDLVPGKVKMLNSVDLHDPYPERAKMDYPMGMITPILPTRT
GTFEILAEILGCVGHPOKMGTVVDTEEDYQAMLAEDQTFQSLSASSETRAVPERVCS
GFPSGIAITEOGTGSALFKEBERECFPGPATTVASAAO"
complement(1457. .1513)
/note="Sm-5 OR SMC04646
REPAT SM-5
predicted by Homology"
/evidence=not_experimental
1634. .3415
/gene="coxN OR SMC01983"
1634. .3415
/gene="coxN OR SMC01983"
/EC_number="1.9.3.1"
/function="small molecule metabolism; energy transfer;
electron transport"
/note="Product confidence : probable
Gene name confidence : probable
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="PROBABLE ALTERNATIVE CYTOCHROME C OXIDASE
POLYPEPTIDE I TRANSMEMBRANE PROTEIN"
/protein_id="CAC47096.1"
/db_xref="GI:15075541"
/transl_table="MMVDVRSIGELAPPEVEDVELYHPHSMTRYVFSQAKITAI
OYMTAIAIGWALVLSMLRLQDGPTEELIDERYOFTIMGMIVYLLTALF
LGPGNVILIPWGVARDWEPYANNISYVIYLAIVLAASFTPGSGAGCTLPV
DAVLSGTGGGRWGLIMLSLIFVIGTGMGINVVVILGARGMTLMRLPLV
GIFATVALLARPALFVACVMLEDRILGTSFPMATYEMKROLOYGCSPLTPOHL
FMFGHPREYVIALPAGFIVSDLISTHARKNIFGRMMWAIYITGLSFTYVAHHNY
VSGMNPYRGFEFATTTLLIATPATAIKVKNVITLMKGNHLTLPMLFALAEIVTVNG
GLTGLFGLNVVDVPLSDTMFVVAHNVAGVAPINVIIGATVHWPKITGRMLNEAM
GOIHFWTFIGAVYATFEFPMVGLIGVPRRYHGLGASVVTISIAELNFIYSVALLV
GNAOIVLENLAMSLRHGRGAGNPWRATVLEMGTPETPPPHGMCNELPVVRYMAYD
YSVGAPEDFITQNPTRGRLSHETYS"
3412. .4110
/gene="coxO OR SMC01984"
3412. .4110
/gene="coxO OR SMC01984"
/EC_number="1.9.3.-"
/function="small molecule metabolism; energy transfer;

```
electron transport"
/note="Product confidence : probable
Gene name confidence : probable
Predicted by Codon_usage
Predicted by Homology
Predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="PROBABLE CYTOCHROME C OXIDASE TRANSMEMBRANE
PROTEIN"
/protein_id="CAC47097.1"
/db_xref="GI:15075542"
/translation="MSIVAFFLAALAIITAMWLAGORTLSRPMLVGHFDRGATRL
PPAKIGLVFLAVVGLAFSLAISAYPMRASSDWMGALPLGLMLNTGILAAGSITTL
WTKEAERRNDENARIGLAGLAGLAFAGLFAWRALSDAGYFLAGNPANSPFYL
TGMGLHIIIGLFFALGRVTAHASQTPPLGNRTLSIELCAIYMHMFLIWLVLAFPG
WASGVIERCOLLT"
4123..4845
/gene="cox OR SMC01985"
4123..4845
/gene="cox OR SMC01985"
/EC_number="1.9.3.-"
/function="small molecule metabolism: energy transfer:
electron transport"
/note="Product confidence : probable
Gene name confidence : probable
Predicted by Codon_usage
Predicted by Homology
Predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="PROBABLE CYTOCHROME C OXIDASE PROTEIN"
/protein_id="CAC47098.1"
/db_xref="GI:15075543"
/translation="MSAPMKKPAATLRRPAGLAGIASPMASDQRTFKDVSCKAMKM
ITLSDTFVFCFLATISAKSSTSVPMNPESEVALTEGTHMPLILAIIMFVLIS
SSGMAAVNGYLRDRRTALMLTLALFGAFAVGMQAFEMSKILAEGRPMGNG
AAQGFSTFTFTGHTGHTVGVFLILIVARKVMGDFETERRGFTSRKREIYEI
TGLYWHFVDLWVTFEAFVFLM"
4857..5216
/gene="SMC01986"
4857..5216
/gene="SMC01986"
/function="miscellaneous: hypothetical/partial homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
Predicted by Codon_usage
Predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="HYPOTHEICAL TRANSMEMBRANE PROTEIN"
/protein_id="CAC47099.1"
/db_xref="GI:15075544"
/translation="MAHAETHPQAHAANTHQHAIKLVLLVWGLIFVLSALSYLD
YRELQGLKMSLIIIFMLKAGLIVAVFMHMAWERTALITAILPLPLIVLVALMVS
ESNYVLTFTLTFEGGGE"
5477..8038
/gene="SMC01987"
5477..8038
/gene="SMC01987"
/function="small molecule metabolism"
/note="Product confidence : putative
Gene name confidence : hypothetical
Predicted by Codon_usage
Predicted by Homology
Predicted by Framed"
/codon_start=1
/transl_table=11
```

Query Match

2.8%; Score 37.8; DB 1; Length 340900;

```
Best Local Similarity 54.7%; Pred. No. 1.5e+02;
Matches 75; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 403 ccgttcgcaaatccctgcgtgatgcgcgcgttcgcattcgtcgaatg 462
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 186372 CCGGTCGATCCACCGGCTTCTCGCGCGGCAAGACGACCTTGAACGGG 186313

QY 463 gacacacatcgctgcgtgcgcacctacgctcatcacaagaagccgcgaagctc 522
|| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 186312 CTAAGGATCCGGCGCTTGCAGACGCGAGTCATCATCAAGATTCGGCATGTCTCC 186253

QY 523 aaacgcgcctgtgtgt 539
| ||| |||
Db 186252 ATGCACACCTCCTCGT 186236

RESULT 37
AY011741 760 bp DNA linear MAM 14-FEB-2001
LOCUS AY011741
DEFINITION Canis familiaris EDG1 (EDG1) gene, partial cds.
ACCESSION AY011741
VERSION AY011741.1 GI:12829947
KEYWORDS
SOURCE
ORGANISM
dog.
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 760)
Murphy,W.J., Eizirik,E., Johnson,W.E., Zhang,Y.P., Ryder,O.A. and
O'Brien,S.J.
Molecular phylogenetics and the origins of placental mammals
NATURE 409 (6820), 614-618 (2001)
2 (bases 1 to 760)
Murphy,W.J., Eizirik,E., Johnson,W.E., Zhang,Y.P., Ryder,O.A. and
O'Brien,S.J.
Direct Submission
Submitted (01-NOV-2000) Laboratory of Genomic Diversity, National
Cancer Institute-FCRCC, P.O. Box B, Bldg. 560, Rm. 11-10,
Frederick, MD 21702, USA
FEATURES
source
1..760
location/Qualifiers
/organism="Canis familiaris"
/db_xref="taxon:9615"
<1..>760
/gene="EDG1"
/product="EDG1"
<1..>760
/gene="EDG1"
<1..>760
/gene="EDG1"
/codon_start=3
/product="EDG1"
/protein_id="AAK02009.1"
/db_xref="GI:12829948"
/translation="VFILICFIIILENIFVLLITMKRFRPMYFIFGNLALSDLLA
GVAATANLLISGATYKLPAPQWPLREGSPFALSASVSLAIAERYITLMKRIH
NGSMSERFELISACWISIVGLPIMGWNCIGATASQVLPVIAKHVILICTTYV
TLIIATVTCRIYSIVRTSRRTFRNISKASSSEKSLALKTIVITLVSVFIAC
WAPLIFLLLVGCKVKTCHILFRREVYFLAVL"

BASE COUNT 127 a 262 c 188 g 183 t

Query Match 2.8%; Score 37.6; DB 4; Length 760;
Best Local Similarity 48.6%; Pred. No. 1.1e+02;
Matches 103; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 1010 gctgggttcgcgcgcgcgaataactcatcagcgcacactctcgccattcc 1069
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 277 GCTACATCAGCATGCTGAAGATGAAGATCAGACGCGGACAAACAGACTTCGCTCTCC 336

QY 1070 taaaaaacaactcttcaagttcacgacagcgctcaagcgcgcgacatggtac 1129
```

Db 337 TGCATCATCAGCGCTGCTGGCTCATCTCCCTGCTGGCGGCTGCTGCCATCATGGGCT 396
QY 1130 cgcctgacactatgagcgagcgttgccttgagacatctgctacatctgcttgccgcg 1189
Db 397 GGAACCTCATCTGGCGGCTGGCGCTGCTCCACCTGCTGCTGCTTACACCAAGCACT 456
QY 1190 attaatcctgcgcgacatcacgacgacgacg 1221
Db 457 ATATCTCTTCTGCAACACCGCTCTTACGCTG 488

RESULT 38
KPUB1143 1731 bp DNA linear BCT 31-MAR-1998
LOCUS Klebsiella pneumoniae heat shock protein (HSP60Kp) gene, complete
DEFINITION cds.
ACCESSION U81143 GI:3002493
VERSION U81143.1
KEYWORDS Klebsiella pneumoniae.
SOURCE Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
ORGANISM Klebsiella
REFERENCE 1 (bases 1 to 1731)
AUTHORS Cancino-Diaz, M.E., Curriel-Quezada, E., Cancino-Diaz, J.C.,
Dominguez-Lopez, M.L. and Garcia-Latorre, E.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-1996) Immunologia, Escuela Nacional de Ciencias
Biologicas, Instituto Politecnico Nacional, Prolongacion de Carpio
y Plan de Ayala s/n, Mexico City 11340, Mexico
Location/Qualifiers
1. 1731
/organism="Klebsiella pneumoniae"
/db_xref="taxon:573"
88. 1731
/gene="HSP60Kp"
88. 1731
/note="HSP60Kp"
/note="groEL-like"
/codon_start=1
/transl_table=11
/product="heat shock protein"
/protein_id="AAC08714.1"
/db_xref="GI:3002494"

BASE COUNT 457 a 437 c 513 g 324 t
ORIGIN

Query Match 2.8%; Score 37.2; DB 1; Length 1731;
Best Local Similarity 45.5%; Pred. No. 1.5e+02;
Matches 132; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

QY 267 agtcgagatgcgcttgaagcaagcaagatgagctgagcagctacgtacgtacgc 326
Db 600 AGGCGTATACCGTTGAAGCGGTACCGGCTGGAAGCAACGACGAGCTGTTGAAGG 659
QY 327 gctggcaaaattgacgacgcaaaaagtgcgcgcaacctgattcaataagcgtatgac 386
Db 660 TATGCAATTGAGACCGGCTGCTGCTCTACTTCAACAAGCGGACACCGGCTGC 719
QY 387 tgcgcttgcacccgctgcttgcagaaataccctgcgcgtagatgcgcgacgttcgcat 446

Db 720 CGTTGAGCTGGAAAGCCGTTTCATCTGCTGCTGACAAAAAATCTCAACATCCGCGA 779
QY 447 ctctgcaatgcatgagatgacaccaatcgcctgcgcgacacctacgcatcatcaaga 506
Db 780 AATGCTCCCGCTTCTGGAAGCCGTAAGCGAAAGCGGCAACCGCTGCTCATCTGCTGA 839
QY 507 agcgccgaagactcaaacgacgctgtgtatgtgacgcgcgtacgcg 556
Db 840 AGACGTTGAAGCGAAGCGCTGCGGACCCCTGCTGTTAACACATGCGCG 889

RESULT 39
AF190427/c 4130 bp DNA linear PLN 25-OCT-2000
LOCUS Colletotrichum lindemuthianum putative GAL4-like transcriptional
DEFINITION activator (CUTAI) gene, complete cds.
ACCESSION AF190427
VERSION AF190427.1 GI:10998377
KEYWORDS Colletotrichum lindemuthianum.
SOURCE Colletotrichum lindemuthianum.
ORGANISM Colletotrichum lindemuthianum.
REFERENCE 1 (bases 1 to 4130)
AUTHORS Sordariomyces Incertae sedis; Phyllochorales; Phyllochoraceae;
Sordariomyces Incertae sedis; Phyllochorales; Phyllochoraceae;
1 (bases 1 to 4130)
Dufresne, M., Perfect, S., Pellier, A.L., Bailey, J.A. and Langin, T.
A GAL4-like protein is involved in the switch between biotrophic
and necrotrophic phases of the infection process of Colletotrichum
lindemuthianum on common bean
Plant Cell 12 (9), 1579-1590 (2000)
11006333
2 (bases 1 to 4130)
REFERENCE Dufresne, M., Perfect, S., Pellier, A.L., Bailey, J.A. and Langin, T.
AUTHORS Direct Submission
TITLE Submitted (27-SEP-1999) Sainsbury Laboratory, John Innes Center,
Colney Lane, Norwich NR4 7UH, United Kingdom
JOURNAL Location/Qualifiers
1. 4130
/organism="Colletotrichum lindemuthianum"
/strain="ups9"
/db_xref="taxon:5458"
join(1801, 1591, 1654, 1673, 1750, 1850, 1902, 2129,
2180, 2794, 2844, 4130)
/gene="CUTAI"
/product="putative GAL4-like transcriptional activator"
801. 4130
/gene="CUTAI"
join(1522, 1591, 1654, 1673, 1750, 1850, 1902, 2129,
2180, 2794, 2844, 4050)
/gene="CUTAI"
/note="belongs to the zinc cluster protein family of
fungal transcriptional activators"
/codon_start=1
/product="putative GAL4-like transcriptional activator"
/protein_id="AAG25917.1"
/db_xref="GI:10998378"

BASE COUNT 894 a 1274 c 1052 g 910 t
ORIGIN

GENE
CDS
801. 4130
/gene="CUTAI"
join(1522, 1591, 1654, 1673, 1750, 1850, 1902, 2129,
2180, 2794, 2844, 4050)
/gene="CUTAI"
/note="belongs to the zinc cluster protein family of
fungal transcriptional activators"
/codon_start=1
/product="putative GAL4-like transcriptional activator"
/protein_id="AAG25917.1"
/db_xref="GI:10998378"

Db 83911 GCTCCAGAACACGCGTGGGGGTGCAGTTCGCCCTGTCCAAAGCGGTTCGCAAGAACT 83970
Qy 1196 tctgtcgagacataccagcagcgagcgttggtggttgcttggaattgacgaagaagacc 1255
Db 83971 TCGTGTGACGACACCGAGATTCCCGACAGCGCGCAACACACGCGTGTGGAACCGGCGC 84030
Qy 1256 tgcgttttgagcgttcgtctgcgcgcggggaatacgaatacgcgcgtgtgtgcgaag 1315
Db 84031 TGACGATGCGCACTACTTCGGGGCGCGCAATTGACGCGACGCTGCTGCGCCACG 84090
Qy 1316 tgcgtg 1321
Db 84091 GCGTGG 84096

RESULT 42
AE008180/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AE008180 10029 bp DNA linear BCT 18-DEC-2001
Agrobacterium tumefaciens str. C58 circular chromosome, section 238
of 254 of the complete sequence.
AE008180 AE007869
AE008180.1 GI:15157875
Agrobacterium tumefaciens str. C58 (Cereon).
Agrobacterium tumefaciens; alpha subdivision; Rhizobiaceae group;
Bacteria; Proteobacteria; Rhizobium.
1 (bases 1 to 10029)
Hinkle,G., Slater,S.C. and Goodner,B.
Complete Genome Sequence of Agrobacterium tumefaciens C58
(Rhizobium radiobacter C58), the Causative Agent of Crown Gall
Disease in Plants
Unpublished
2 (bases 1 to 10029)
Hinkle,G., Slater,S.C. and Goodner,B.
Direct Submission
Submitted (14-AUG-2001) Cereon Genomics, 45 Sidney Street,
Cambridge, MA 02139, USA
FEATURES
source
1. 10029
Location/Qualifiers
/organism="Agrobacterium tumefaciens str. C58 (Cereon)"
/strain="C58"
/db_xref="taxon:181661"
157. 615
/gene="AGR_C_4831"
157. 615
/gene="AGR_C_4831"
/note="Marr family regulator"
/codon_start=1
/transl_table=1
/product="AGR_C_4831p"
/protein_id="AAK8387.1"
/db_xref="GI:15157876"
/translation="MGTRKREKDSLAFLLNSGARLLNNAFERRISEAGLGTPEGARAL
LTVAAVDSKOSDIARIGIEPTICAYLDKLSLDLIERODPDORRSKRLVLTNS
TDLLEAVROEIDELVRHATRLNDEVALFRSLQTLKDNDPEPGQSC"
621. 1817
/gene="AGR_C_4834"
621. 1817
/gene="AGR_C_4834"
/note="(AE006190) Bcr"
/codon_start=1
/transl_table=1
/product="AGR_C_4834p"
/protein_id="AAK8388.1"
/db_xref="GI:15157877"
/translation="MPEPIMSTRKTAIIIGLITLALPISMAIYTPAMELTRVATN
ESAKISLSTYFAGFALAQILAGPASAFGRRRSLIYVFLGGGLAIFATPTEVL
LVARLGGIGASVGMVRAIVRDQFTGAESSLNNLGLIMAVPALGPTIGGLSLA
ATGKSVFELFMAGIAAIFSVVLEFRETIVPKSLIPRRLISAYGNLTPEPFLIA
AVLGGSIGALYAQATMLTFILINTVGMTPTAFGIMGMOTGAYFGESIALYAPRL
GRRSYMLGICFSGTGLLILLSVFLIPSPFSIMGPAVAATVGTALILTPSVATAA

PEPHIAGSASAMMGFIOMGSGFAGCTAASILGSLPTGTGFIYIPVMEFTAAVASYIGPRI
VSRSKT"
complement(1872. 3254)
/gene="AGR_C_4836"
complement(1872. 3254)
/gene="AGR_C_4836"
/note="argininosuccinate synthase PA3525 (imported) -
Pseudomonas aeruginosa (strain PA01)"
/codon_start=1
/transl_table=1
/product="AGR_C_4836p"
/protein_id="AAK8389.1"
/db_xref="GI:15157878"
/translation="MAGHPSSCIPINCFAVYLTERKREFTHGLHCKSAEIPRKNPCPD
AVIMODIIMALPKDVKVYLAVSGGIDTSLIKWLOTLEGAEVVETPAADICGGELE
PARKKEMGIKEIYEDVREDFRPDMFRNRYVEGYLLGTSTARPLISHLI
ELAKTGAGALAHGATGKNDQVRELSATLANPDIKTIAPRWMTKSRDLNFAE
OHQIIPAKDKGEAPSVDAANLHSSSEGVLEDPAAVEPYVHMRITISPAADPKAT
VTKVGRKDKDACSINGVSPATLNLNMGRENGIGRLDVENREVMKSRGYET
PGCTILAAHRAIESITLDRGAHLKDLMPRYAELIYIGWFSPEMLOALIDKSO
EHVEGEVTLKLYGNVMVGRSEKSLYSALVTFEDDHGAYDCKDAAGFIKLNALRL
RTLAKRNLNK"
3327. 4265
/gene="AGR_C_4837"
3327. 4265
/gene="AGR_C_4837"
/note="probable oxidoreductase - Streptomyces coelicolor"
/codon_start=1
/transl_table=1
/product="AGR_C_4837p"
/protein_id="AAK8390.1"
/db_xref="GI:15157879"
/translation="MKTVEATQVQATLQTAGNSRNDAREKPMPTIHPAEKNVAVI
TGAASIGIGLAARKFASFGMSVYLVDLDEKRLAAHADIILNADGGAQIYATPDVS
KIDELALERAVIQRGRVHYLMNNGIOPCSAIFGQAMKEKIIIGNIMGVNYSRV
RGQGITTHREPSLITWTGSKQGITTPGDDAIVNSGVATFETALHELNITNSAV
SAHLIIPGVFTGLTAHGREPKQAGMGTQVDFMTESIGRGDFYILCPDGEVDRT
DEKRILMAQDVIQVNRPLSRMHTESAAFTSLNK"
4338. 5045
/gene="AGR_C_4838"
4338. 5045
/gene="AGR_C_4838"
/note="conserved hypothetical protein - Deinococcus
radiodurans (strain RJ)"
/codon_start=1
/transl_table=1
/product="AGR_C_4838p"
/protein_id="AAK8391.1"
/db_xref="GI:15157880"
/translation="MAETRNRYKQPHSPDFTGAMDPVSLPTLIIAFTIALIIVAP
GPDMTLWISRLREGRAGFNTLVGTNIGITVATMDVAFGVAALIVASPTAFMLKTG
GAAYLVWLAIQAIRKSGDFVWVKSTEKQASLSKALNLIWNLLPKYIIFEMTEL
FOFVSATDPHVTKGLIFLGIMSIIVALPIGIGIVVADILSAMLQNRNRYLRGLDYTI
AGVSLFAVKIPEFTQTR"
complement(5084. 5878)
/gene="AGR_C_4841"
complement(5084. 5878)
/gene="AGR_C_4841"
/note="probable ATP-binding component of ABC transporter
PA388 (imported) - Pseudomonas aeruginosa (strain PA01)"
/codon_start=1
/transl_table=1
/product="AGR_C_4841p"
/protein_id="AAK8392.1"
/db_xref="GI:15157881"
/translation="MISLSDIOWFGRGPTPILOKQALAKIDLTDIEDGSFYVYISNGAG
KSTILIGVAGVLPAGKVMIGGADVTRKPTASRAGVAVFQDPLAGSGALITEN
LALASRGERGLSSALDGRGDFRFRALNGLNKLKMDLLSGQROAVSYV
MATLSGSVILLIDETALIDPGMAEFVMLTFRKVSERKLTITLAVTHSMQADYGR
TIMHAGRIYVLDGSDSRKLDLEVEDLIEMFRTIRGQTLDDDELIC"
complement(5875. 6756)
/gene="AGR_C_4842"
complement(5875. 6756)

/gene="AGR_C.4842"
/note="probable ABC transporter, permease protein VC1102
(imported) - Vibrio cholerae (group O1 strain N16961)"
/codon_start=1
/transl_table=11
/product="AGR_C.4842p"
/db_xref="GI:15157882"
/protein_id="AK88393.1"
/translation="MSQIAFWGAVELGVFAVAIGVIAFRVLDPEPDLTVDSFPIG
AAVAALVITAGLAWAVATAIAMVAGAAGAVTATLVKFKTLMILASITLMIAFSVN
LRYMGKPVALLNODTMVGFYGLSEYLVRLPVGVLVIAIVLWRFLESDACIA
MRATGANARARAQGVKTGNQIYGLMALNALVALGALPAGTNGFADVSGVTIVV
GLAAVITGETEGARGLIALIGCVGSIATRIALQIALSSDVLGLKASDLNVTANV
VALILRLRREGATS"

gene
complement(6863..7861)
/gene="AGR_C.4844"
complement(6863..7861)
/gene="AGR_C.4844"
/note="hypothetical protein PA8336 (imported) -
Pseudomonas aeruginosa (strain PAO1)"
/codon_start=1
/transl_table=11
/product="AGR_C.4844p"
/protein_id="AK88394.1"
/db_xref="GI:15157883"
/translation="MFLERRKOGTNRIRISMRAPFLIALAASATVLPKAKENVVAATA
IYEHPLDARQGVAKLAAGYKKEGKGNILKELYESAGNGCTAAQIARQVADAPAVI
VPSWPSAAVVAATRDIPVTFVAVSDPVGAOLIKSEKRGAVTGSDMLPGEHIA
LIKEISPNKASIGFIYNSAEANSATLALAKENDKGLVLSVATKSKEVOCATGA
LVGKADVIYVPDNTVSAFEAVAGVASEKIPLYAADTVSARGAVALGAFNPDVG
KQGAIVVRLKKEKPGDIPATVAVGTDLVINKVAKVGVPEPESVTKRATKATVD"
complement(7925..9265)
/gene="AGR_C.4846"
complement(7925..9265)
/gene="AGR_C.4846"
/note="conserved hypothetical protein NMB1308 (imported) -
Neisseria meningitidis (group B strain MD58)"
/codon_start=1
/transl_table=11
/product="AGR_C.4846p"
/protein_id="AK88395.1"
/db_xref="GI:15157884"
/translation="MRKCSAFPLDMQTFRLFDPRVYVCMERWNVKSMVQAPAGL
AKAPILANDLSLGSKPSLGLTFRKEGALAEIVPQKVKMRPVSQLMWLIVRGVS
DDEDMNTNVAKELEKLEKAAFTIARPELVEQIINDGRKMLMRPPGARGAPVELET
YIDEGRGTCISSQVGCSTCSFCHGTGRLNLAEEILSQLLRARLGPFPFG
STPVGAIVSEGRKGVNIMVMGEPLEYNFENHTALLITDGGELISKRVTLS
GVVPEIFRTGDEIGVMALISLHVRDLRDLINKYELKEIEACRYVPGVSNAR
RIFEYVMLKDVNDSLDDAKMLVOLLKGPAPKINLIPNPWPGTNYCCSMAOIEKFA
DRIQNGYASPRITPRGRDILAACGLKSESEMRKTERLAFEMATANGADD"
complement(9364..9897)
/gene="AGR_C.4849"
complement(9364..9897)
/gene="AGR_C.4849"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="AGR_C.4849p"
/protein_id="AK88396.1"
/db_xref="GI:15157885"
/translation="MEDNPNAGSMFVRSVATAAAILTSVGLASASPPRIQOFNAM
GAYSYSNGSTVCYLSTIPSKSPANSDHGDIFIVSORGONISYEPQAMVGTLLQ
GSKVNTIDKNKFNVFTKDKAAVENVAAEPAALVAMKGGKSMYKAVSRGATATSS
YSLSGISAFAKOIESCK"

BASE COUNT 2072 a 3048 c 2888 g 2021 t
ORIGIN

Query Match 2.88; Score 37; DB 1; Length 10029;
Best Local Similarity 48.88; Pred. No. 1.9e+02;
Matches 100; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 343 agcgaaagtgcgcgaacctgattcaatcagcgtatgactgcgctcgacccgt 402

Db 8104 ATCGAAAAATTCGCGCATTCATCATCAGGCGGTTATGCGTCCCATCGCAGCGC 8045
QY 403 ccgttcagcaaatccctccgtagatgcgagccgttcgcatctgtcaatgcagt 462
Db 8044 CGCGCGCGCGCATTCCTCGCGCCGCTCGCGCATTTGAAGTGGAGATTCGAACCGATGGCG 7985
QY 463 gacaccaatccgctgctgcgcgacccctacgcatcatcaagaacgcgcgaagattc 522
Db 7984 AAGACCGGAAGACTGCGCTTTTGAGCGCATGATGATGCCAACGCGGCAGATGACTGA 7925
QY 523 aaacgcgcgcctgtgtatgagcc 547
Db 7924 ATTAAGCCGCGTTTGCGCATTAAC 7900

RESULT 43
AE009214/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

1 (bases 1 to 10359)
Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E.,
Chen, Y., Woo, L., Kitajima, J.P., Okura, V.K., Almeida Jr., N.F.,
Zhou, Y., Boye, Sr., D., Chapman, P., Clendenning, J., Deatherage, G.,
Gillet, W., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M.,
McClelland, E., Palmeri, A., Raymond, C., Rouse, G.,
Saenphimmachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I.,
Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M.,
Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C.,
Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V.,
and Nester, E.W.
The Genome of the Natural Genetic Engineer Agrobacterium
tumefaciens C58
Science 294 (5550), 2317-2323 (2001)

2 (bases 1 to 10359)
Wood, D.W., Setubal, J.C., Kaul, R., Monks, D., Chen, L., Wood, G.E.,
Chen, Y., Woo, L., Kitajima, J.P., Okura, V.K., Almeida Jr., N.F.,
Zhou, Y., Boye, Sr., D., Chapman, P., Clendenning, J., Deatherage, G.,
Gillet, W., Grant, C., Guenther, D., Kutyavin, T., Levy, R., Li, M.,
McClelland, E., Palmeri, A., Raymond, C., Rouse, G.,
Saenphimmachak, C., Wu, Z., Gordon, D., Eisen, J.A., Paulsen, I.,
Karp, P., Romero, P., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M.,
Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C.,
Zhao, Z., Dolan, M., Tingey, S.V., Tomb, J., Gordon, M.P., Olson, M.V.,
and Nester, E.W.
Direct Submission
Submitted (27-SEP-2001) Department of Microbiology, University of
Washington, 1959 NE Pacific Ave, Box 357242, Seattle, WA
98195-7242, USA

FEATURES
source
Location/Qualifiers
1..10359
/organism="Agrobacterium tumefaciens str. C58 (U.
Washington)"
/strain="C58"
/db_xref="taxon:180835"
complement(84..1037)
/gene="Atu2672"
complement(84..1037)
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"

gene
CDS
/protein_id="AAL43653.1"
/db_xref="gi:17741177"
/translation="MRAFLIALAAASAIYLPKAKENVYAVYAIYEHPLAARDGVK
AALAEYKKEGNEKELYESAGNGCTAAQIAKQVGVGAPVIVYISTPSAAVYAT
RDIPYFTAVSDPVCAQILKSMERFGVNTGSDMLPYGEHLALIKETSPKNSIGFI
IVSAEAVAGVASEAKIPIYAADDSVARGAVNALGFNVYFDVQKQGAIVVRLKGEK
PGDIPATVAVGDIYINKKVAEKVGVPEPESVTKRATYID"
complement(1146..2318)
/gene="Atu2673"
complement(1146..2318)
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/transl_table=1
/codon_start=1
/product="conserved hypothetical protein"
/protein_id="AAL43654.1"
/db_xref="gi:17741178"
/translation="MSGRSLGLTREMEGALAEICYPQKQYKRVSOQMLLYVRC
VSDPNMTNVAKEKELKAAFTIARPELVEQISNDGTRKMLRPPPGAGRPVEIE
TVYIPEGRGTLCTISSQVCSLTCSECHGTQRLVRLNTAEILSQLLARLDLPEP
DGSIPVAGVYVSEGRKVNINVMGMEPLYNFEHVKTALLATDGDGLSKRYVLS
TSGVYPEIFRTGDEIGVMLAISLHVRDLDMALVPINKKYPLEKIEACRYPCVSN
ARITFEYVMLKDVNDLSDAKMLVOLKGVPAKINLIPENPMPGNOCSEMAQIEK
FADFNOGVASPIRTPRGROIACGOLKSESEMRKTERLAFAMIANHCAD"
2537..3034
/gene="Atu2674"
2537..3034
/gene="Atu2674"
/note="putative; ORF located using Glimmer"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAL43655.1"
/db_xref="gi:17741179"
/translation="MAGRPANHKSDKASGLLAALDESSRNARQIRIGIRSAAG
NRLVGHRLAHLHGDESRLLFSGVNPISVIGEDVLYVDRHIDPLGALLQIANHRL
RLVGVYLAARTLDEDEDAIVVACRFLLGGDGKNAVDAVARFVIGTPIELDARG
RLRC"
complement(3386..3766)
/gene="Atu2675"
complement(3386..3766)
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAL43656.1"
/db_xref="gi:17741180"
/translation="MDVYKIGELILPGETDKORENENYRAKFWPKLRKVMTKVPEAR
DAAYVYCAIDRDPILRAKGIILALAFYIMPIDAVPMLAVGFTDIAVITATAM
IRAHKMEHYEADMLKROQEG"
complement(3818..4123)
/gene="Atu2676"
complement(3818..4123)
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="4a-hydroxytetrahydrobiopterin dehydratase"
/protein_id="AAL43657.1"
/db_xref="gi:17741181"
/translation="MKYRLEPDALRQALSKLEGSRLREDGAIVRSKFSSEFAEAFG
FMASALAEKLNHPEMNSYRVKVCITLTHDSQVTERDFLLAEAMOKAAGRGN"
4183..4662
/gene="Atu2677"
4183..4662
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"

gene
CDS
/codon_start=1
/transl_table=1
/product="protein tyrosine phosphatase"
/protein_id="AAL43658.1"
/db_xref="gi:17741182"
/translation="MKHTVLEFVCMGNCIRSLABGVTLDLADGEGVSQRTVDSAGT
GGMHQCDDPDRSIAMARRHGDISRQARQVTRADFEAFDLILAMDENNLNLQLS
PEKYRKHILFENDYAGHENYPPDPYGGAEDEGLDLYVNMMLLAGCRSLLEKMLDRAS"
complement(4596..5189)
/gene="Atu2678"
/note="Atu2678"
complement(4596..5189)
/gene="Atu2679"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/transl_table=1
/codon_start=1
/product="2'-5' RNA ligase"
/protein_id="AAL43659.1"
/db_xref="gi:17741183"
/translation="MRLETALEIPRNAAMSLIRGLGPGARWIDENYHITLRFIG
DIDGRTADEVDRLDIRIEPFOQLMTGIGSGSKRPHSIWAGVSPSPMALQAEVE
RICQRIQHPDPRKPMPTIARLRSSRYDVVHYLSGRGNRTSPPTGVREVLMSSK
ESVGGAPYVEAPFLEHARSSISNDIHPAKSML"
complement(5258..5908)
/gene="Atu2679"
complement(5258..5908)
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/transl_table=1
/codon_start=1
/product="esterase"
/protein_id="AAL43660.1"
/db_xref="gi:17741184"
/translation="MREKALHFITVILAGAFSSAVAQOETLQVLGIGSLMAGYQ
LPTDTSYQLEPAALKARGINVAITNAGVSGDPSGGLARASVPDGTGVYILEGA
NDALGALPEQTEKMDITLIISGFORNIATLVGIMAPNMGDDYAKRNPILPRLAE
KYNLEPYFEFLGVVTDALKEDKMHPYTKGVAVVVEKSLPAGESFIKTGAQRK"
5999..6709
/gene="Atu2680"
5999..6709
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="ABC transporter, nucleotide binding/ATPase
protein"
/protein_id="AAL43661.1"
/db_xref="gi:17741185"
/translation="MTKSIKLRKADLTGLNAAASVHLKNIDLSIDEGEAVGIVGPS
GSGKSTLLMVLAGERLDSGEIVADTOLHKGEDLADFRGNIGIYFOSPHLAINM
TALENVAVLELANTPNPEIKARELVAAGLERNNHYPGQSGEQQVATAPALAP
SPAVALADEPTGNTDPTGKQIADLLFAKQABRGMTWLVTHDPSLAARCSQIVRS
GEIEGDSARPOKARAVSA"
6706..9267
/gene="Atu2681"
6706..9267
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAL43662.1"
/db_xref="gi:17741186"
/translation="MIGSIPSPFTNIRNAARLARREIRGLRGFYIFLACTISLGTGAI
AAVNSVRSATTEGOSILAGVRFVRLRRRATAEBSYLDGIGTVAVSTGLRSM
ARLADGSOITLFEVATIDGAYPLRFPADPBRPLLAGSGDIPYAAVAPLILRL
GIKVDDELILNAKRLRTGVYDEPDALSDGCPAPRIYVSDALPASGLIOTGSIVE
HAYKITLSDPARATMTDANKAFPOAGWISITSDRAAPSLTENYNRFSQFLITVLGLT

Query Match	Best Local Similarity	4.88:	Score 37:	DB 1:	Length 10359:
Matches 100:	Conservative	0:	Mismatches 105:	Indels	Gaps
QY 343	agcgaataagtcgcgcgaacctgatcaatcagctatgacgcttcgcacacct				
Db 1325	ATCGAAAAATTCGCCGACCTTCATCATGACGCGGGTATGCGTCGCATCGACGCGC				
QY 403	ccgttcagcaaatccctccgcgttagatgacgagccgttcgcacatcttcgtaatgcgatg				
Db 1265	CGCGGCCCGCAGCATTCCTCCGCCCTCGCGCCAGTTGAAGTCGGAAATGMAACGCAATGCGC				
QY 463	gacacaaatccgcttgcgtcgcgacacctacggtlcatcatcaagaagacgcgcgaagaattcc				
Db 1205	AAGACCGAANAACGTGGCTTTTGAAGCGCATGATGCGCAACGACGCGCAGATACGTGA				
QY 523	aaacgcgcgcgttggtatgtagac				
Db 1145	ATMAACCGCGTTTGCGCCATAAAC				

RESULT	LOCUS	DEFINITION	ACCESSION VERSION	KEYWORDS	SOURCE ORGANISM	REFERENCE AUTHORS	TITLE	JOURNAL MEDLINE REFERENCE	AUTHORS
AE002070/c	10514 bp	DNA	linear	BCT 22-NOV-1999					
AE002070	Deinococcus radiodurans RI section 207 of the complete chromosome 1.								
AE002070	AE000513								
AE002070.1	GI:6460209								
	Deinococcus radiodurans.								
	Deinococcus radiodurans Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcaceae; Deinococcus.								
	1 (bases 1 to 10514)								
	White,O., Eisen,J.A., Heidelberg,J.F., Hickey,E.K., Peterson,J.D., Dodson,R.J., Haft,D.H., Gwinn,M.L., Nelson,W.C., Richardson,D.L., Moffat,K.S., Qin,H., Jiang,L., Pamphile,W., Crosby,M., Shen,M., Vamathevan,J.J., Lam,P., McDonald,L., Uitterback,T., Zalewski,C., Makarova,K.S., Aravind,L., Daly,M.J., Fraser,C.M. et al.								
	Genome sequence of the radioresistant bacterium Deinococcus radiodurans RI								
	Science 286 (5444), 1571-1577 (1999)								
	20036896								
	2 (bases 1 to 10514)								
	White,O., Eisen,J.A., Heidelberg,J.F., Hickey,E.K., Peterson,J.D., Dodson,R.J., Haft,D.H., Gwinn,M.L., Nelson,W.C., Richardson,D.L., Moffat,K.S., Qin,H., Jiang,L., Pamphile,W., Crosby,M., Shen,M., Vamathevan,J.J., Lam,P., McDonald,L., Uitterback,T., Zalewski,C., Makarova,K.S., Aravind,L., Daly,M.J., Minton,K.W., Fleischmann,R.D., Ketchum,K.A., Nelson,K.E., Salzberg,S., Smith,H.O., Venter,J.C. and Fraser,C.M.								
	Direct Submission								
	Submitted (08-NOV-1999) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA								
	Location/Qualifiers								
	1..10514								
	/organism="Deinococcus radiodurans"								
	/strain="RI"								
	/db_xref="taxon:1299"								
	/chromosome="1"								
	93..2717								
	/gene="DR2392"								
	93..2717								
	/note="identified by Glimmer2; putative"								
	/codon_start=1								
	/transl_table=11								
	/product="hypothetical protein"								
	/protein_id="AAFI1946.1"								
	/db_xref="GI:6460217"								
	/translation="MYVSDDTICNNODGVRGVORLYICYPLKGGTNVLRYLWDELISVSLVGSTAOIEVKRLALQADDSSEVRQLRDLDFGKLTSPTITVRNNEPRLPDMAEFGEGYATRTLEGVGLTEIQPSQSPSLKLDVILDEGVKYTRYEVTPDRSRSGYGVLS								

ATGLDGFPELQDVTYDNRARAGYSGPLGACGLYYAAAKDGLPTDNRNLINTNYGSAAL
TESVLDQVQIDVAFSTYDHPQRAEYRSLSTIDVLPGEANTLALATSKINPRYSCV
ALVPEDRVANNRPLPEEGRRLRLDGDYSGLSGETLEVTLESTRGKELQREKLVKRW
YVLPDRSCVVFENRARDLDELNERVYVASYRLDNPPLANRTAYGAQIYTTGTINSSV
GYAAVYLDGRTFTYGVKADYORGPKVASKGLAYSGGVOASADLSTVLRQNDHNEVRX
QOASVEGLAPAPGLVYSGRYTGRLTPLNVAADAEYRDMSPRASDVENANNAHNG
VTARVYHLDPFTYQAGYSYAFGDRSGGLGVSGYREPTSYDVHTQPIQDLPDT
TDISRFPIRDKDTLGFNDKLTWCGGAAALTLDTMGVNAVAYGVELPARGSGNRA
RFGATVTLPLNDRNLALNRGNVLYDVCGGAGEGATGADINKTDRLSATYQDVSYRG
DFGCVFVLRGCVTGSDVDELTLNAGGTAEFGCGRGERLELYAYRGRSLSLGYGRY
VAGSLAGKGPPELSPGVSAAVRYLPMNAVYGVGSEYRDLDDRSPFLQGLGSSYLTDR
LAVGMAARLDTLPQASQTLTFCGYGVGYSRPLPGWITLACGYNFRBEGRQSSYTRRG
PYLRDLDTLDETLGGPQR"

gene
287. .3701
/gene="DR2393"
287. .3701
/gene="DR2393"
/note="similar to GB:Pyro_h percent identity: 53.56;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAFI1939.1"
/db_xref="GI:6460210"
/translation="MTSHLRGILLLVLTVAIWGSTFAYVKEGLGALLAPPVYLLAMRSIT
GALVYLPLAALRRTPAPTPVTVTQADGTSLSWSDGVALGLMLTAGGTQIALQTTNAR
AAFFALSSVVLVPLVLTQRRRMAPLMAALPLAVGLALLSMEGAGWSDGVALAR
CAYVAGTGLILEKTLASRHAALPTLAAQVLSLVAVMGMALLSGAPLMPPOAMAAPLE
LVGVATVGTGTLDTLQGRHVSAAEASLYALPEVSAALSFALIGERVAGALGSA
LVYLVATVSSRGETEPYARVLDSSQD"
3894. .5702
/gene="DR2394"
3894. .5702
/gene="DR2394"
/note="similar to PID:1177251 SP:P50864 PID:1644212
GB:AU09126 percent identity: 51.54; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="N-acetylglucosyl-L-alanine amidase"
/protein_id="AAFI1940.1"
/db_xref="GI:6460211"
/translation="MKLPAMLLSSALLIGSMAAQOTDPEFLRABAYQKPYVYRGSAAP
PSAASFSTPSAQTPLQAQAKETTFGSPRISGTEATKLVDELPRKMTYTLVPTFGGL
RLDVGVSVAASSLRLALETEYRANNGQIIMQTPFLSLSDMGATEATIANGNV
LIVELGSLRGAGVAGGRVLAAPADPRAQSOAPVPYDTPAASALTSHELPCGDS
VKQAPRPPAPLAPLPGADTAPRSALVQSGTSSGATVTPTRIGNPGVTRTGL
PPGASRYVEPTGLRYVELLGVGLDANTNEONTSAEYRMRTPPSAEVNLSTVTS
PTTSSGNRAOVLPLNERSRLVLDISPAANALPLSPRKLVAAPPMTASGAALV
LSLRASVPRVYITDPHGKGDPGATGATEKDYTLDAARVRLVAAGVADVLTRE
TDLREKPNKNDLQAKRMKGTQGMFYSIVNNATEAVARSLAGGIIWMNPINALSY
OLATVLDQDVTGCAFPGRGIKGYRSLSVENSRPALVIEIGTSHPVQGNLADNN
YLDRAMKIASGIRPALMTGVSAAKPKPAAVGGP"
CDS
599
/gene="DR2395"
599
/gene="DR2395"
/note="similar to GP:332762 percent identity: 64.01;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=1
/product="Na⁺/H⁺ antiporter, putative"
/protein_id="AAFI1941.1"
/db_xref="GI:6460212"
/translation="MTWTCRLPEAPRYTDMLTAPALLSTVNAFLNENRFPYPTT
VGVTLGALSSMLLALDALDGLTGRGMAAGLLELTDFTDVLNGILSLTLFAALDGL
DARQLRRRFSTLAVSTSLSTFLTGLFAAYGVFGVLGVLPMWLSLFEALISPD
RYVALDILKRAVPRRIETTLJAGESLEFNDGVGVIFVILVAGMAGISHSYGVSVAL
ELFGEALGAGMAGALGALGIMGLREFNDGVADEVLITLAVMGVAAASLTPGPD
AMVVGATVSGRDVAREGEERTREHESMEPTDQVNLITLPAFGIDVLITETPGPD
VAGLLVVASLARRISVALRPPILRRHEDYGYTVRLLTTCGLRGCTAISLYGLDQ
TPYRHLVATVYITVLFSLIAVGGLTLMPLVHKAAVAPSDSAGPDSAGGQEDGENA
GROLSSEET"

gene		7192.. .7788 /gene="DR2396"
CDS		7192.. .7788 /note="similar to GB:A1123456 percent identity: 52.33; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="transcriptional repressor, Tetr family" /protein_id="AAFI1942.1" /db_xref="GI:6460213"
gene		/translation="MPEPRRRLSAEVRREQLILEVASOLFIERGFENVKADIAEHLOT SPNITVTYPTSTEAIIIDILLERMGDMLOEAVRYADNSLFPDDMLRALERELL LLYSGGGPPRSRORRVLDSDDOHTYNLPENVGTANPDLLIVQRTLVSTAHELLH PEYDROIQOLIMELVSARKERPDIQAKHSRA" complement(7839..8645) /gene="DR2397"
CDS		complement(7839..8645) /gene="DR2397" /note="similar to GB:U00096 PID:1789101 PID:882637 percent identity: 61.28; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="survival protein SurP" /protein_id="AAFI1943.1" /db_xref="GI:6460214"
gene		/translation="MTVPDPAPPAASPSARPPVLVANDGITFAPGKALGAISEWAD VVYPADVDEOSAVGHGITTIRPLRFKHTAAGFGDIPAYVDGTPADCVYLGLGR PDLVSGINIGNLGEDLTSHGVAAAI EGLTLGPSIAFSQFAVEAGEYDFPSAY ASRLAAEVCCRGSLPRLVLNVNPFPSRGVRYEEVGJHRWDSVYTRDDPECDYHW VAGVSVAHDGHDEQDTGYGAOVAQGISVSPVRLDIFARDLIGELTQALPPL" 8712.. .9314 /gene="DR2398"
CDS		8712.. .9314 /gene="DR2398" /note="similar to GP:2815343 percent identity: 58.15; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAFI1944.1" /db_xref="GI:6460215"
gene		/translation="MTVIRTLEDLESLEGEYGABSLQKETDYLSHYROWLEAPPAV LATSGGGGLDASPRGDPAFLRVYIDERTLLPERGNRNRIIDLRLISDPRGILFEI GVNETLRVNGTGARIITLTDLLSFADDGKLPKCYLETVDFVFQPMARSVMWT GSSAEVPTPGAMLAALTREVDGADYDAALPQRQRDLY" complement(9345..10322) /gene="DR2399"
CDS		complement(9345..10322) /gene="DR2399" /note="similar to GB:A1123456 percent identity: 54.93; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAFI1945.1" /db_xref="GI:6460216"
gene		/translation="MSAPILPCGALYPEPHSCQGPROMSGTELGMADGARGVGRARPH RCGTPILAGRESAOPAGSRRAGAGPLDPRRAVRAHLIDGQAFHBEQDETTEVYS ITAFPRTDSPHCVALTHSRSGRALTGMVNLGSEYVNTAOEYVELIDONTADGFL AADALEAOVDALSEOEIFRSPFNVEVPFDKLHGQARRITLDAREASALLGROAST AAEOYRVPODSFTRVGSRLDCLRESLSLIDLHLAIOGORAMEYRTLITSYVFL PLTFLAGVGMKFOHMDELATRGTAARWLFTLLIGLLTIWFKRGMW"
BASE COUNT	1771 a 3750 c 3474 g 1519 t	
ORIGIN		

Db	6450	GCGCTGCTGGGGGCCCATGTGGGTTCTTATGCTAGCTGC	CGGAATTCGACACGACCCGCTCGAG	6391
QY	1078	aactcttcaagtttacgacacgcgcgtcaacggcgagcgcgcgcattgltacgcgatcgcgc	1137	
Db	6390	GTGCTCATTACACCCTTGCGGCTGTGCTCGGGGGATCACTGAGGGGCTCGCACTGGGCGTC	6331	
QY	1138	actatgagcgcgctaataatgcgtttgacaacctgcctacccttgctcttttcgcgatttaac	1197	
Db	6330	AGCGGCCGCTGAGGCGATGTGTGTCGCGGTGTCGATCTCGCGGGGCGGACCTGTGGCC	6271	
QY	1198	gtcgcgcgtaccacagacagcgcgcgcagcgtttgttgttcgttgatttgaagaagaacatc	1257	
Db	6270	TTCGGGAGAAGAACCCCGGAACACATCGAAGATTTCTGGAAACGACCGATCAGTGCTC	6211	
RESULT 45				
LOCUS	M2355			
DEFINITION	M2355	3855 bp	DNA	linear UNA 04-AUG-1993
		Figure 2. Nucleotide sequence of the 3,854-bp HindIII fragment encompassing the hnp locus and the deduced amino acid sequences of putative polypeptides encoded on the sense strand.		
ACCESSION	M2355			
VERSION	M2355.1	GI:340948		
KEYWORDS				
SOURCE	.			
ORGANISM	unclassified.			
	unclassified.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 3855)			
TITLE	Mukhopadhyay,P., Williams,J. and Mills,D.			
	Molecular Analysis of a Pathogenicity Locus in Pseudomonas syringae pv. syringae			
JOURNAL	J. Bacteriol. 170, 5479-5488 (1988)			
MEDLINE	89053865			
FEATURES				
source	Location/Qualifiers			
	1..3855			
	/organism="unclassified"			
	/db_xref="taxon:12908"			
BASE COUNT	741 a 1178 c 1181 g 755 t			
ORIGIN				

[illegible]